Band Width and Transmission Performance

By C. B. FELDMAN and W. R. BENNETT

In modern communication theory band width plays an important role as a transmission parameter. The authors discuss the significance of signal band width and frequency occupancy in relation to other transmission factors such as power, noise, interference, and overall performance for certain specific multiplex systems under assumed operating conditions. The intent of the paper is to show how such problems may be attacked rather than to find an unequivocally best system.

The scope of the paper is described by the following table of Headings and Captions.

List of Frequently Used Symbols

B = radio signal band width in megacycles. (Not to be confused with frequency occupancy).
b = base or radix of PCM system.
β = peak-to-peak frequency swing of FM systems in megacycles.
F_b = width of baseband (video band) in megacycles.
f_r = repetition or sampling frequency in megacycles.
K = load rating factor (amplitude ratio).
log = logarithm to base 10.
ln = logarithm to base e.
N = number of channels in a multiplex system.
n = number of digits in a PCM system or number of spans in a multirepeater system.
P = wanted carrier amplitude.
P_n = mean fluctuation noise power per megacycle.
Q = interfering carrier amplitude.
S = span length in miles.
U = band spacing factor.

I. Introduction

CARRIER systems for tbe transmission of many telephone channels on a single metallic circuit have grown to be very important in the telephone network. Since the development of the coaxial cahle system in which 480 channels are transmitted in a 2-mc baseband, advances in high frequency techniques, including the war-accelerated microwave art, have inspired efforts to utilize the broad band capahilities of high transmission frequencies. Some of the efforts bave related to the wave-guide conductor but mainly they relate to radio relay transmission. As a consequence of these efforts a considerable number of new multiplex methods for use at microwave frequencies have been devised. All of these methods employ handwidth more liberally than the 4 kc per channel rate associated with single sideband carrier systems, in return for which various transmission advantages are ohtained. Theoretically, transmission advantages can he sacrificed to permit handwidth reduction but tbe transmission requirements then become very severe. Bandwidth as a transmission parameter has grown to a prominent position in modern communication theory as set forth hy Shannon et al.[1, 2, 3]

The liheral use of bandwidth, employed in an effective way, operates to permit higher noise and distortion within a system and, in the case of radio relay systems, operates to permit higher interfering signals from other radio systems. When all the frequency space necessary to avoid mutual inter-

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Sys. Tech. Jl.*, Vol. 27, pp. 379–423, 623–654, July–Oct. 1948.
[2] B. M. Oliver, J. R. Pierce and C. E. Shannon, "The Philodophy of PCM," *Proc. I. R. E.*, Vol. 36 (1948), pp. 1324–1331.
[3] C. E. Shannon, "Communication in the Presence of Noise," *Proc. I.R.E.*, Vol. 37 (1949), pp. 10–21.

ference between systems in a congested area is taken into account, certain wide-band methods, less vulnerable to interference, may be as or more efficient in the use of frequency space than other narrower band multiplex methods.

The principal purpose of this paper is to examine, for various systems, the relations governing the exchange between frequency space and transmission advantages.

It will be shown that the preferred multiplex method depends in part upon:

1. The grade of facility required; low-grade and high-grade channels lead to different preferences. These preferences also are influenced by the length of circuit.

2. The nature of the transmission obstacle over which advantage is sought. These obstacles may be: (a) intrasystem distortion (phase distortion, overload distortion, etc.) and noise; (b) intersystem interference as between similar radio systems or between different types of radio systems, operating on the same frequency.

Other factors beside the transmission considerations discussed here are likely to be involved in a practical multiplex application; hence the system preferences arrived at in this study may not be the controlling factors in practice.

Before a detailed analysis is undertaken, it may be helpful to examine and comment upon the chart shown in Fig. 1. All of the multiplex methods shown here have been studied sufficiently to permit their approximate evaluation with the aid of some theoretical considerations and subject to certain qualifications as pointed out from time to time. Variations and combinations of these are possible,[4] some of which will be discussed later.

In addition to the two general classifications of frequency and time division there is a third type based on carrier phase discrimination. A familiar example is the quadrature carrier system,[5] which is capable of yielding two channels for each double sideband width. In another form[6] each of N channels is modulated simultaneously on $N/2$ carriers with a different set of carrier phases provided for each channel. Time division multiplex may be regarded as a kind of phase discrimination in which the signal is modulated on harmonic carriers so phased as to balance out except during the channel sampling time intervals. In true phase discrimination,

[4] A comprehensive listing and discussion of various combinations will be found in a recent paper by V. D. Landon, "Theoretical Analysis of Various Systems of Multiplex Transmission" *R.C.A. Review*, vol. IX, numbers 2 and 3, June–Sept. 1948, pp. 287–351, 438–482.

[5] H. Nyquist, "Certain Topics in Telegraph Transmission Theory," *A.I.E.E. Trans.*, April, 1928, pp. 617–644.

[6] W. R. Bennett, "Time Division Multiplex Systems," *Bell Sys. Tech. Jl.* Vol. 20, pp. 199–221, April, 1941.

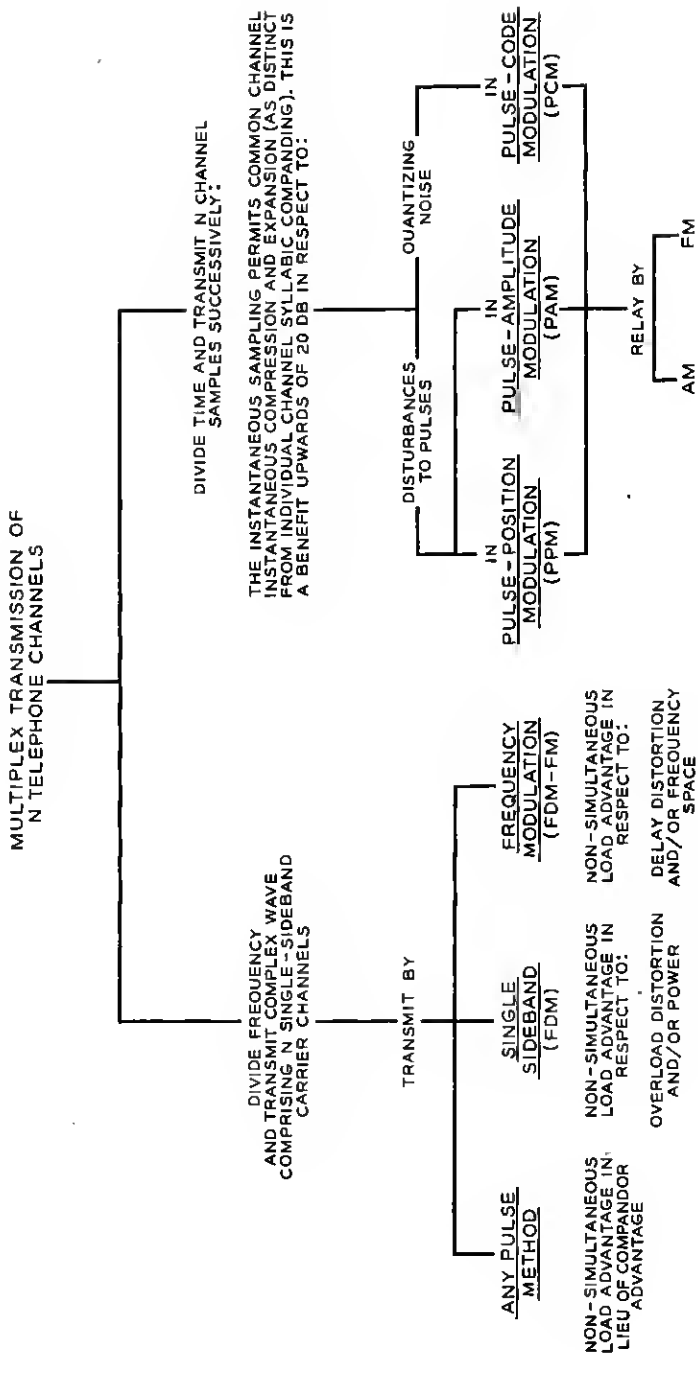


Fig. 1—Outline of multiplex transmission methods.

however, there need be no separation of channels in either time or frequency, and a homodyne detection process is required at the receiver for channel selection. The necessary precision of instrumentation seems in general more difficult to achieve than with either frequency or time division, and only minor prospects appear for exchange of handwidth for transmission advantage.

Of the systems tabulated, the frequency division method (FDM) with single-sideband suppressed-carrier transmission is the only method in which bandwidth cannot be traded for some transmission advantage.[7] This system will be used as a standard of comparison. The PAM method, with transmission by AM pulses, can trade upon bandwidth only as a means for reducing interchannel crosstalk. In the other pulse systems, as well as all systems using FM, handwidth may be expended to gain advantage over noise, intersystem interference, and, generally speaking, intrasystem distortion and noise.

Non-simultaneous Load Advantage in FDM

The *non-simultaneous load advantage* pertaining to frequency division multiplex refers to the fact that the channel sidebands rarely add to an instantaneous value even approaching the value N times the peak value for one channel. This means that the required peak capacity of a relay system transmitting the N channels increases slowly with N. Current toll transmission practice provides for relative power capacity roughly as follows.[8]

Table I
Non-Simultaneous Multiplex Load Advantage

N	Required Relative Power Capacity	Advantage
1	0 db	0
10	+6 db	20 − 6 = 14 db
100	+9 db	40 − 9 = 31 db
500	+13 db	54 − 13 = 41 db
1000	+16 db	60 − 16 = 44 db

To emphasize the strikingly large non-simultaneous load advantage statistically obtainahle with conversational speech we may examine Table I and note, for instance, that the capacity of a 1000-channel system is completely

[7] We have in mind here a system such as Type K or L in which a minimum separation of adjacent channels in frequency is used. It is true that by spreading the channels far apart in frequency, a reduction in cross-modulation falling in individual channels could be obtained, but the resulting amount of improvement is minor compared with that offered by a corresponding band increase in the other systems.

[8] B. D. Holbrook and J. T. Dixon, "Load Rating Theory for Multichannel Amplifiers," *Bell Sys. Tech. Jl.*, Vol. 18, pp. 624–644, October, 1939. The values in the table come from curve C, Fig. 7, taking the single channel sine wave power capacity as +9.5 dbm.

used up by peak instantaneous voltage when 994 channels are disconnected and 6 carry full-load tones.

If a group of carrier channels in frequency-division multiplex were translated to microwave frequencies, the overload distortion affecting the transmission would be predominantly of the third-order class. To a first approximation the third order distortion follows a cube law and may be predicted from the single-frequency compression. We assume here that the power capacity of the repeater is the output at which the single frequency compression occurring through the complete system does not exceed 1 db.[9] This criterion applies roughly to systems of several hundred channels capacity, and to present transmission standards.

Instantaneous Companding Advantage in Time Division

In time-division systems, as ordinarily understood and known in the current literature, each channel successively is provided with its full-load capacity, and thus a non-simultaneous load advantage does not accrue. However, because of the sampling process, instantaneous compression may be applied at the transmitting terminal before noise and distortion are encountered; when complementary expansion is applied at the receiving terminal the noise is suppressed. The expanded samples derived at the receiving terminal then bear an improved relation to noise, particularly in the case of weak samples. Such an instantaneous companding process applied without sampling to a continuous speech wave requires a greatly increased transmission band between compressor and expandor but, in a time division system, no more bandwidth is needed to transmit the speech *samples* after they have been compressed than before. An instantaneous compandor currently being used experimentally to handle 12 channels in time division has the noise performance characteristics shown[10] in Fig. 2. It is shown as applied to a telephone system in which the channel noise power (unweighted) would be 45 db down from the power of a sine wave which employs the full load capacity provided for the "loudest talker". Abrupt overloading is assumed to take place when peak amplitudes exceed that of the full-load tone. The location, at -7.5 on the load scale, for the power representing the very loud talker (one in a thousand) conforms approximately to current practice. The speech volumes, referred to the point of zero db transmission level, are shown for the sake of completeness.

[9] In a multi-repeater system the compression accumulates. This means that each repeater must be restricted to operate approximately 10 log n db below the 1 db compression point of one repeater. (n denotes the number of repeaters.) See Section VI.

[10] Use of the same curve to represent the performance with tone or speech implies an independence of wave form which is not rigorously valid. Calculations based on speech-like signals have indicated that the curve for tone loading is a good approximation when average power is used as the criterion in the manner shown.

The compression and expansion result in a uniform improvement of 26 db for weak signals including the "very weak talker" and a lesser improvement for stronger signals. The noise power in the absence of speech is 71 db

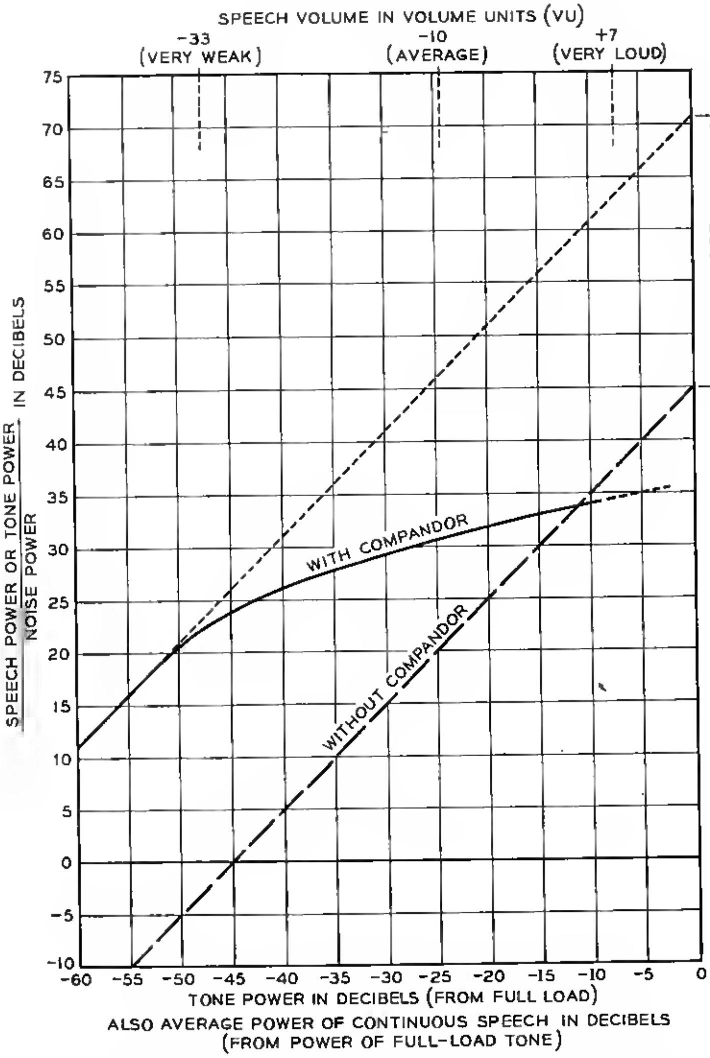


Fig. 2—Performance of an experimental instantaneous compandor.

below the power of the full load sine wave which the channel is designed to handle. The performance is substantially equivalent for telephone purposes to a 71-db circuit without companding, in spite of the fact that for all except the very weak talkers the average noise is greater than with the 71-db circuit.

The noise is increased (over the 71-db value) by the compandor only in the presence of speech and then only in proportion (roughly) to the amplitude, and so becomes masked by the speech. The masking is sufficient to make impairment of medium and loud speech imperceptible provided that the ratio of speech power to noise power is greater than about 22 db. Under these conditions we are justified in defining the equivalent signal-to-noise ratio in terms of the low level noise.

For compandors with a more drastic characteristic, yielding more low-level improvement, the high-level noise increase is enhanced and the limit to this enhancement is controlled by the "uncompanded" signal-to-noise ratio (the ratio without companding.) Thus the amount of low-level improvement that is permissible from the standpoint of high-level performance is determined by the uncompanded signal-to-noise ratio. Experiments have shown that the permissible low-level improvement increases several db for each db increase in uncompanded signal-to-noise ratio. Another way of putting it is that the value of the equivalent signal-to-noise ratio in the speech channel determines the amount of compandor advantage which may be invoked to attain that ratio, and that the permissible compandor contribution increases nearly as fast as the equivalent signal-to-noise ratio. The uncompanded signal-to-noise ratio is thus required to increase only slightly.

For the 45 db uncompanded signal-to-noise ratio of Fig. 2 the compandor could have been designed to yield more than the 26 db low-level improvement shown without impairing high-level performance. In the time-division systems of message grade with which we will deal later, a 22 db compandor advantage is assumed.[11]

In the quantized systems included in the PCM heading the instantaneous compandor advantage applies to the granularity, or quantizing, noise in the same way as to the common kinds of noise which plague other systems. The compandor of Fig. 2 was actually used in an experimental PCM system.[12] A discussion of quantizing noise appears in Appendix I and a more comprehensive treatment appeared in the Bell System Technical Journal recently.[13]

In transmitting frequency divided groups of channels by pulse methods[14]

[11] This is the maximum compandor advantage permissible for a circuit equivalent to 57 db signal-to-noise ratio. We will use this figure in connection with power requirements for circuits whose signal-to-noise ratio is intended to be equivalent to 60 db but since we presume that interference or crosstalk may be present in an amount equal to noise and since the compandor acts on interference as on noise we must protect against high level impairment on the basis that the noise is 3 db greater.

[12] L. A. Meacham and E. Peterson, "An Experimental Multichannel Pulse Code Modulation System of Toll Quality", *Bell Sys. Tech. Jl.*, Vol. 27, pp. 1–43, Jan. 1948.

[13] W. R. Bennett, "Spectra of Quantized Signals," *Bell Sys. Tech. Jl.* Vol. 27, pp. 446–472, July, 1948.

[14] If the group occupies a frequency range extending from zero to F_b, the minimum sampling rate is well known to be $2F_b$. If the group range does not start at zero frequency the minimum sampling rate is not twice the highest frequency of the group but lies between two and four times the width of the band depending on the location of the band. This matter is treated in Appendix V.

the instantaneous compandor advantage is suhstantially zero because, at full system load, companding actually increases the total noise. In time division the noise is increased at full load hy companding hut, as discussed earlier, this is permissible hecause full load occurs only with loud talkers who mask the noise. In a frequency divided group transmitted hy pulse methods nearly full load may he produced when a numher of loud talkers are momentarily active; the weak talkers then enjoy no improvement due to companding but may, on the contrary, suffer some degradation.

Non-simultaneous Load Advantage in Pulse Transmission

Transmission of a frequency-divided group hy pulse methods does, however, permit the realization of a portion of the non-simultaneous load ad-

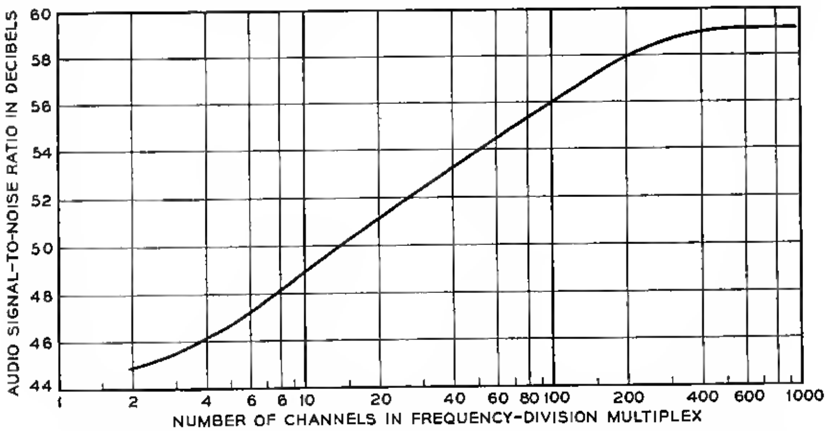


Fig. 3—Quantizing noise in each channel when PCM (128 equal steps) is applied to an FDM group.

vantage in lieu of the instantaneous compandor advantage realizable in time division. A pulse system, designed to carry N channels in time-division with a certain full load signal-to-noise ratio (without companding) may also he used to carry N channels in frequency division. These N channels in frequency division may he treated as a single channel N times wider than the time-divided channels and sampled N times faster. The ratio of the full-load signal in this wide hand to the noise in the wide hand turns out to he the same as the corresponding ratio in each of the narrow, time-divided channels. This fact makes the transmission of large groups of frequency-divided channels advantageous compared to small groups. For instance, in a 100-channel frequency-divided group, a single channel would have availahle a total load capacity which is 9 dh less than that for the multiplex group. This comes from Tahle I, which shows that 100 channels require

9 db more range than one channel. This makes the signal-to-noise ratio in a channel 9 db lower than if all of the entire load capacity were devoted to that channel. However, in the 100-channel group a single channel receives only 1% of the noise power in the entire band, so a 20 db improvement accrues on this score. The net improvement is $20 - 9 = 11$ db. Applied to a 128 step PCM system in which the full-load signal-to-noise ratio is 45 db[15] (7 digits binary PCM), the full-load signal-to-noise ratio in one channel of a 100-channel group thus becomes $45 + 11 = 56$ db. With smaller groups than 100 channels the signal-to-noise ratio falls to 45 db while for larger it reaches 59 db, as shown in Fig. 3. Better results than these are obtainable with time division and instantaneous companding, as shown in Fig. 2, but these results may have significance in relation to the transmission of television by a pulse method, such as PCM. If a 128-step system were used, a large frequency-divided group of telephone channels filling the television band could be substituted for television when desired.

A more powerful application of the non-simultaneous load advantage in time division will be discussed later in Section V.

Signal Band Width and Frequency Occupancy

We define *signal band width* as the width of the signal spectrum (or more realistically as that portion of the signal spectrum which must be preserved in order to make the signal sufficiently undistorted). *Frequency occupancy* is greater than signal bandwidth in two respects:

First, the frequency range accepted by the receiving filter at the end of each span must be greater than the signal band for reasons of filter imperfection. In all of the pulse or FM systems it would be advantageous from the circuit point of view to make the receiving filter much wider in order thereby to reduce the phase distortion over a small central frequency range occupied by the signal band. The assigned frequency space must include the entire band accepted by the receiving filter. Our comparisons will assume that the filters make use of an appropriate amount of refinement to conserve frequency space.

Second, frequency occupancy must include the multiplication of assignments made necessary to avoid interference between converging or intersecting radio relay routes, between the two directions of a single route, or between a main route and a spur.

Our procedure in evaluating these systems will be to plot for each system certain curves relating power, signal bandwidth and channel signal-to-noise ratio or signal-to-interference ratio for various associated transmission

[15] Appendix I shows that the quantizing noise power at the minimum sampling frequency is the same for wide and narrow signal bands. This illustrates the general principle used here.

conditions. From these curves and other pertinent data we will prepare tables which sbow the significant frequency occupancy for various radio relay conditions. Sucb tables will be made for two grades of transmission facilities and for the extremes of signal bandwidth, one corresponding to minimum power and the other to minimum bandwidtb. The minimum power condition prevails when the bandwidth has been increased, and the power reduced, to the point where any further increase of bandwidth would require an increase of power to prevent noise from "breaking" either the pulse slicer or the FM limiters.[16] The minimum bandwidth condition occurs when any further band limitation operates to impair the signal too much, assuming that the power is ample to override noise.

Regeneration and Re-shaping

Two distinct classes of relay operation exist, one applying to the quantized systems (PCM) and tbe other applying to non-quantized systems. When the transmitted signal is intended to convey a continuous range of values (amplitude, time or frequency) noise and distortion accumulate as the signal progresses from repeater to repeater over a relay route. If, however, a range of values is represented by a discrete (quantized) value, a signal may suffer displacement within the boundaries of that range without altering the information conveyed by the signal. If, tberefore, in one span of tbe relay route the displacement is confined to those boundaries the signal may be regenerated and re-transmitted as good as new. No accumulation of noise and distortion need occur, therefore, as the signal traverses span after span. The most common application of regenerative repeatering is in printing telegraphy where the signal is either a mark or space and, if correctly determined, may be re-transmitted afresh.

In all of the non-quantized systems the repeaters must have low distortion so that a signal may be conveyed through a large number of them (say 133 for a 4000-mile circuit made up of 30-mile spans) witbout too much mutilation. In spite of good repeater design a signal passing through such a large number of repeaters will accumulate considerable noise, interference from other systems, and distortion characterizing the repeater design limitations. In non-quantized systems there is no escaping accumulation of this sort. In pulse systems, for instance, phase distortion, common in flat band repeaters, may result in tails and the like, while cumulative frequency discrimination (band narrowing), characterizing simple forms of linear phase repeaters, results in cumulative broadening of the pulses. In the former case the tails

[16] In this connection it is of interest to mention that if the objective were a *very low grade* circuit the power required to prevent breaking might be higher than that required by a method having no improvement threshold, and no power saving could be accomplished by the bandwidth exchange principle. For circuits of telephone grade this situation does not occur.

may eventually grow large enough to break the slicer (if the system employs such a device) while in the latter case the reduced pulse slope and the spreading out of time bounds may also bring about transmission disaster. In both cases these growing distortions successively reduce the margin that it is necessary to provide for noise and interference. To circumvent such effects, the pulses may be reshaped at all or some repeaters. Reshaping consists of measuring the information conveyed by the pulse (in the time or amplitude dimension) and sending out a new pulse of standard shape possessing that measured characteristic in time or amplitude. This process is distinctly different from regeneration as practiced in quantized systems; in general, reshaping can only be counted upon to confine the rate of accumulation of noise, interference and crosstalk to that of power addition from span to span.

In FM systems any distortion which results in amplitude "modulation" of the FM wave may be treated with limiting at each repeater to prevent such amplitude variation from accumulating and breaking the limiter. Like pulse reshaping, this measure does not stop the accumulation of disturbance to the intelligence. Certain kinds of distortion may be combated by double FM.[17]

Reshaping (or, in the case of FM, limiting) may be employed to conserve power in the systems having an improvement threshold. Without reshaping, the minimum repeater power is the marginal[18] value for the total noise accumulated from all spans. If reshaping is practiced at each repeater the power need be marginal for the noise from only one span. More bandwidth must be used, then, in exchange for the lower power; and, while this in turn increases the marginal power, the result is a net power saving. Tables II and III of Section III illustrate this point and Section VIII illustrates its application to metallic circuits.

The Radio Repeater

Repeaters for relaying television signals must achieve low distortion and we will take a current design and assume that such a repeater represents a basis for discussing the transmission of multiplex telephony by non-quantizing methods. This repeater employs, in the two-way application, four antennas and two frequencies as shown in Fig. 4. It is proposed to transmit 5-mc video television signals by FM in bands spaced 40 mc. The repeater employs double detection and the band separation is effected mainly by the

[17] Leland E. Thompson, "A Microwave Relay System," *Proc. I.R.E.*, Vol. 34, December, 1946, pp. 936–942.

[18] By marginal power is meant that power which just safely exceeds the improvement threshold power. For a given noise level, minimum power is achieved when the bandwidth improvement factor yields the required signal-to-noise ratio in the channel with the power that is marginal for that bandwidth.

selectivity following conversion to intermediate frequency. Microwave receiving filters afford enough selectivity to divert alternate bands into their correct frequency-converting units without disturbing the other bands; and microwave combining filters serve in the transmitting side of the repeater to

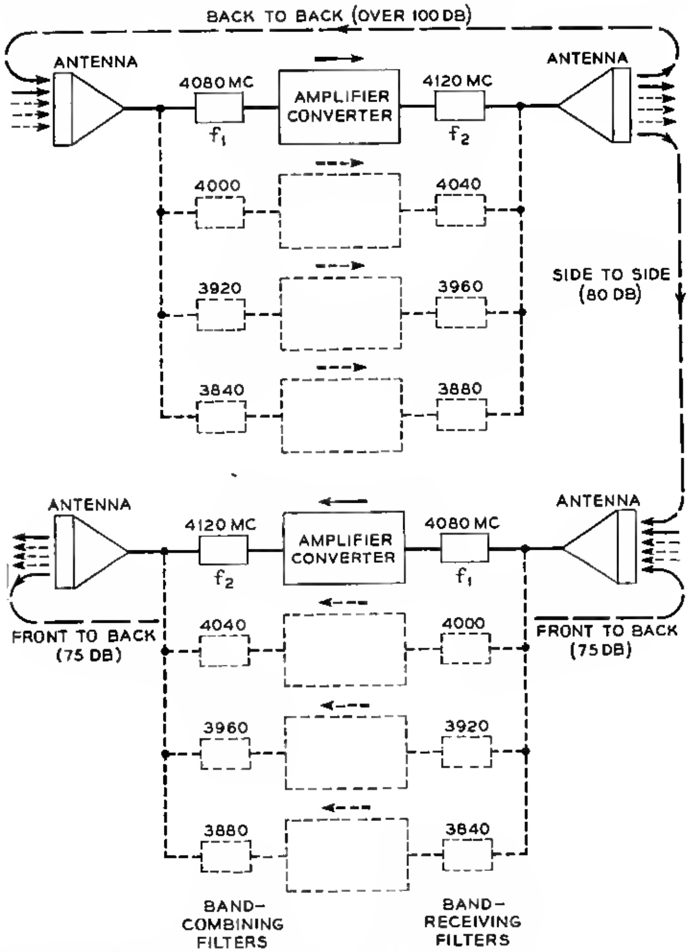


Fig. 4—Arrangement of two-way two-frequency repeater of television type showing spacing of bands and antenna discrimination.

bring the bands into the common antenna with small loss and small mutual disturbance.[19] The combining and separating processes are made easier by the interleaving of transmitting with receiving frequencies. Interleaving

[19] W. D. Lewis and L. C. Tillotson, "A Non-reflecting Branching Filter for Microwaves," *Bell Sys. Tech. Jl.*, Vol. 27, pp. 83–95, January, 1948.

In all of the systems in which bandwidth may be exchanged for tolerance to interference we restrict, in our comparison tables, the minimum bandwidths to those which provide at least the 44 db tolerance demanded in the two-frequency plan.

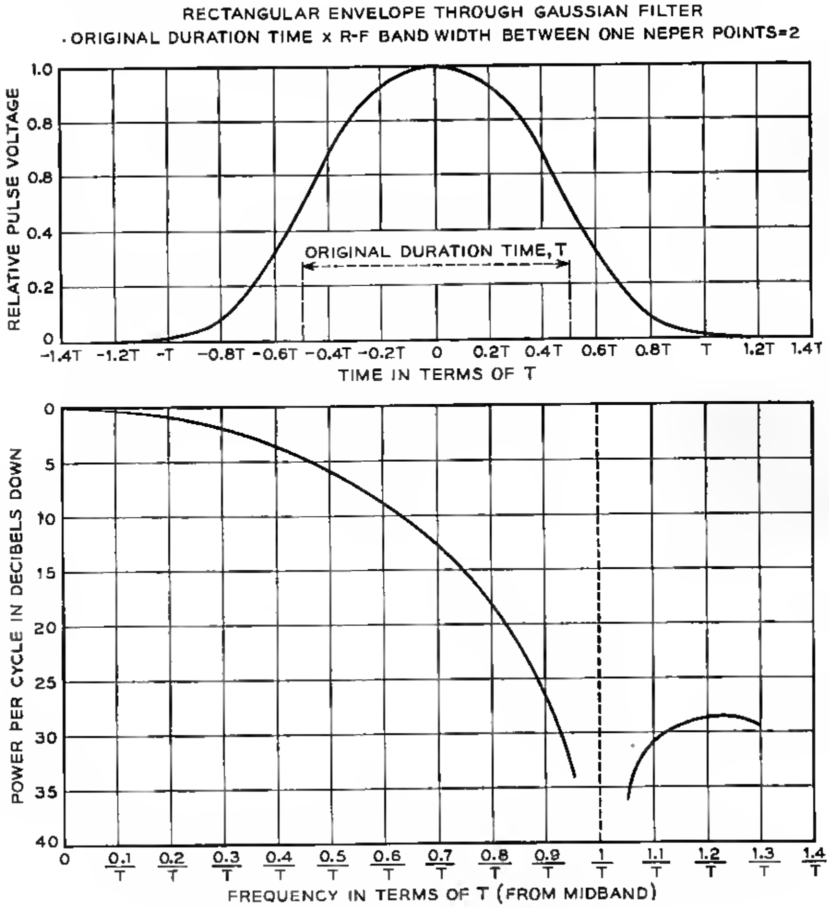


Fig. 6—Basic pulse shape (approximately sinusoidal) and its spectrum.

II. Band Width Characteristics

The type of pulse assumed in the various pulse transmission systems is shown on Fig. 6. As pulse 4 of Fig. 26, it is further discussed in Section VII. The spectral density or distribution of energy vs. frequency associated with such a pulse is also shown. It is evident from this curve that omission of frequencies beyond a baseband width $1/T$ can result in distortion or tails of

only a few per cent of the pulse height. We define signal bandwidth for the pulse systems studied here as $1/T$, or $2/T$ in the *r-f* medium, i.e., double sideband is assumed in all of the AM pulse cases. In assuming double sideband, we bow to the obvious difficulty of dealing circuit-wise with single sideband and its pulse demodulation problem.

In the FM systems we define the radio signal bandwidth as the peak-to-peak frequency swing, β, plus two times the baseband width, $2F_b$.

We shall consider individually the following types of systems where the meaning of the symbols is explained in Fig. 1.

1. PPM-AM	3. PAM-AM	5. PCM-AM	7. FDM
2. PPM-FM	4. PAM-FM	6. PCM-FM	8. FDM-FM

The source of disturbance may be either fluctuation noise, a constant-frequency interfering wave (CW), or a similar but independent system operating on the same frequency allocation. CW interference may fall anywhere within the radio signal band. Interference from echoes, which is a special case of similar system interference, is not treated. In certain cases echoes such as might be produced by multiple reflections in waveguide connections to the top of radio towers may be more detrimental than independent system interference of the same amplitude. We assume that such echoes are suppressed sufficiently by good design.

Our first set of curves, Figs. 9–20, exhibits quantitatively the audio signal-to-noise and audio signal-to-interference ratios which can be obtained with increased radio bandwidth in the various systems. Audio signal is taken to be the power of a test tone which fully loads one channel. Audio noise is expressed as the total noise power in the channel. Audio interference is expressed as the power of all of the extraneous frequencies produced in a channel by the assumed interfering signal. The term "radio bandwidth" is intended to mean double-sideband width and does not imply that the transmission is necessarily by radio. Two of the systems, PAM-AM and FDM, are omitted from this study because, as has been pointed out earlier, they do not provide a significant basis of exchange of bandwidth for suppression of noise and interference. The other systems possess this trading property in varying degree as illustrated by the curves. The FDM and PAM-AM systems are entered in Table IV and discussed under Section III. For comparison with the following curves it may be of interest to note here that for 1000 4-kc message channels in FDM, the bandwidth (single sideband) is 4 mc., and the received power for a 60-db audio signal-to-noise ratio must be -77 dbw.[20] This is the power in a sine wave which employs

[20] Throughout this paper we shall use the abbreviation "dbw" for power expressed in decibels relative to one watt.

the full load capacity in accordance with Table I, at a point where the noise power is −189 dbw per cycle of bandwidth (15 db noise figure, NF).

In most of these curves, plotted for 1000 message channels, the bandwidth scale runs to hundreds of megacycles. We do not mean to imply that the microwave transmission medium can be relied upon to transmit faithfully such wide-band signals or that circuit techniques for producing them are available. As suggested by Fig. 4, the 1000-channel system might be divided into several groups of fewer channels to avoid frequency selective transmission difficulties or circuit limitations. The total frequency occupancy is not altered by such a division, while the required power per group is reduced in proportion to the number of channels.[21]

Curves are shown of audio signal-to-noise ratio as a function of radio signal bandwidth at constant power and at marginal power. Audio signal-to-interference ratios are plotted against radio bandwidth for marginal ratio of radio signal power to interfering signal power. By "marginal power", we mean the radio signal power which just safely exceeds the threshold below which noise or interference causes system failure. In the case of fluctuation noise, any further increment of bandwidth from this point is untenable without an increase in radio signal power. Points on the marginal power curves show as abscissa the bandwidth at which minimum radio power is required to obtain the audio signal-to-noise ratio given by the ordinate. In calculating these curves, we have specified the marginal condition as occurring when the peak disturbance is actually 3 db below the theoretical value which just breaks the system. These relations are shown graphically in Fig. 7. We have in this paper followed the accepted practice of ignoring all fluctuation noise peaks exceeding the rms voltage by more than 12 db. *Radio signal power* is taken as the power averaged over a cycle of the high frequency in the FM wave, or, in the AM pulse case, over a cycle of the high frequency when the pulse is maximum. A curve is included in Fig. 9 showing marginal AM radio pulse power values for various bandwidths of fluctuation noise and a similar curve for FM is shown in Fig. 13. A noise figure of 15 db[22] is assumed for the receiver. We have taken the noise bandwidth as equal to the signal bandwidth throughout. This equality cannot be quite attained in actual systems because of the departure of physical filters from ideal characteristics. In practice an allowance for frequency instability would also have to be included.

The relation of the PPM pulse to channel allotment time is shown in Fig. 8. Pulses in channels adjacent in time can just touch when full load signals are impressed on each. The slicer operates at half the pulse height which, for the assumed pulse shape, is also the point of maximum slope. The width of the

[21] These statements are not exactly true for FDM and FDM-FM, where multiplex load rating is used in the design.
[22] This means that the noise power is 189 db below a watt per cycle of bandwidth.

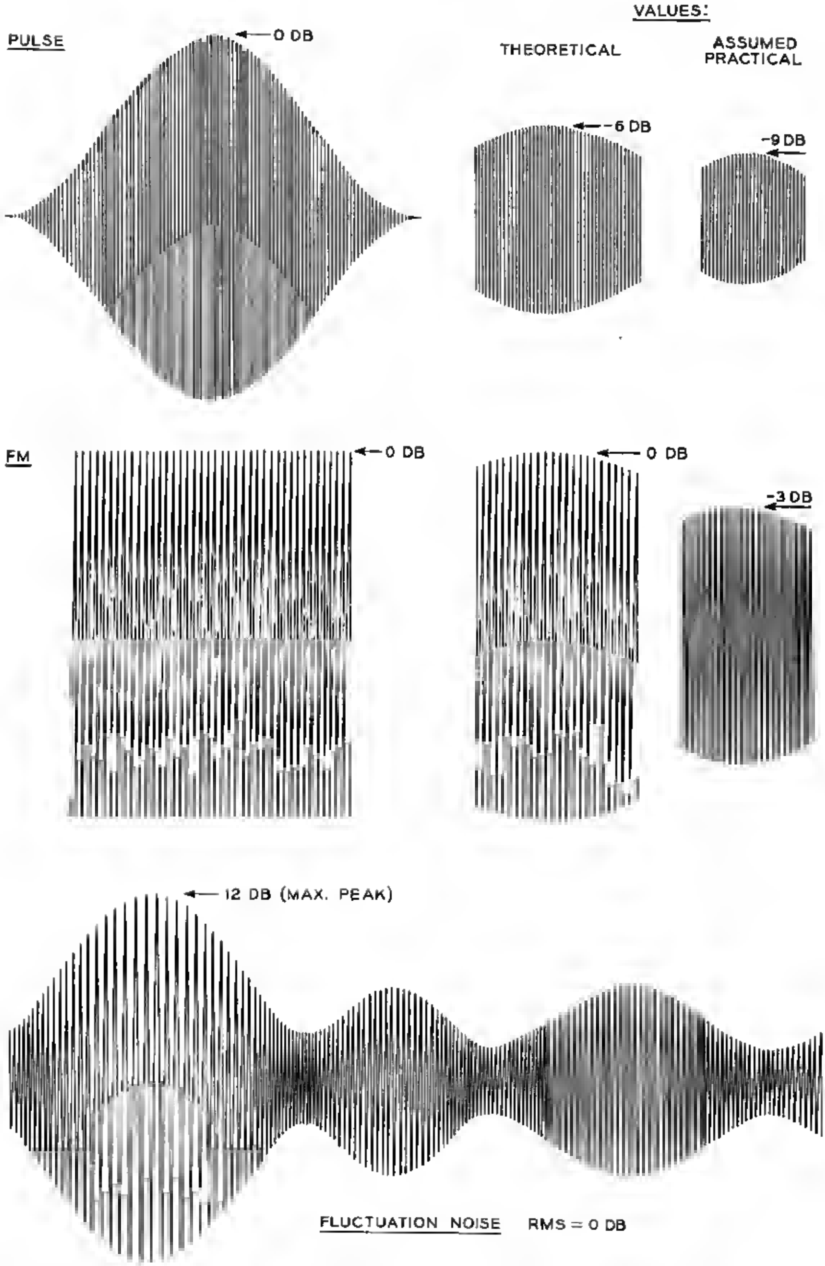


Fig. 7—Marginal condition in reception of AM pulses and an FM wave in presence of noise. Pulse case applies to PCM only if binary.

pulse is inversely proportional to the signal bandwidth. The time available for modulating the pulse position is equal to the channel time minus the pulse duration. The combination of these factors leads to the PPM "slicer

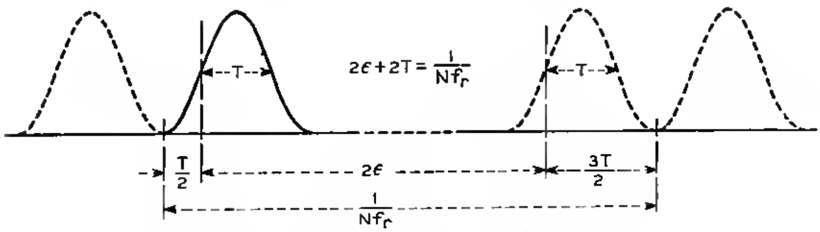


Fig. 8—Time allotments in pulse position modulation.

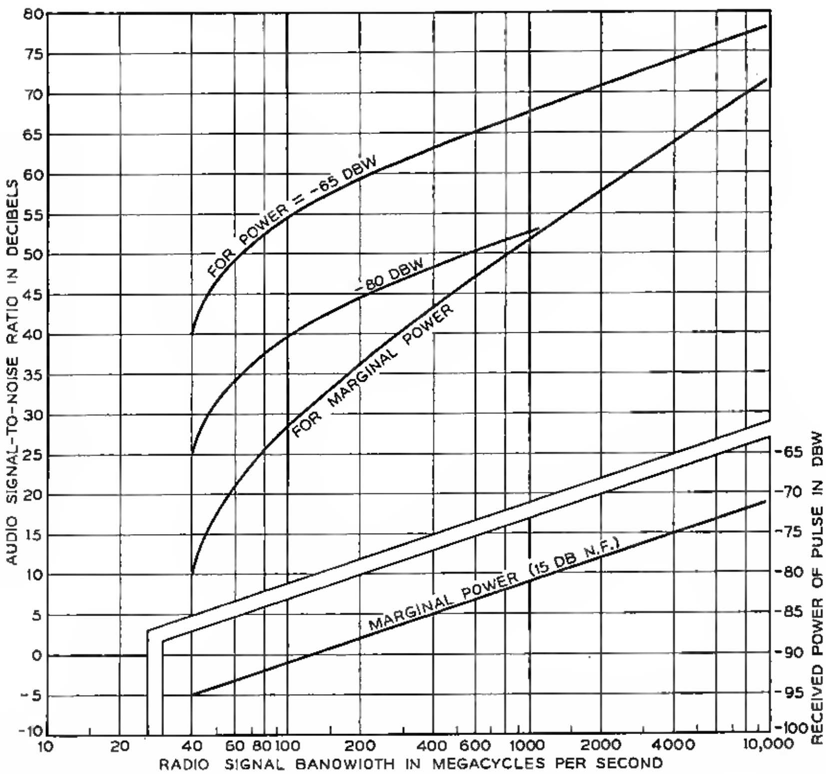


Fig. 9—PPM-AM; performance with respect to fluctuation noise. Relations between bandwidth, power, and audio signal-to-noise ratio for 1000 4-kc channels.

advantage" which, when applied to the *r-f* pulse-to-noise ratio, gives the full load audio tone-to-noise ratio in each channel. Details of this calculation and others pertaining to various pulse systems are included in Appendix IV.

FIG. 9—PPM-AM, FLUCTUATION NOISE

The curves of Fig. 9 were computed from the slicer advantage derived in Appendix IV. The asymptotic slope of the constant power curves of 3 db per octave of bandwidth reflects the 6 db advantage due to the two-fold greater pulse slope (slicer advantage) diminished by the 3 db increase of noise accepted by the two-fold wider band. In the marginal power case

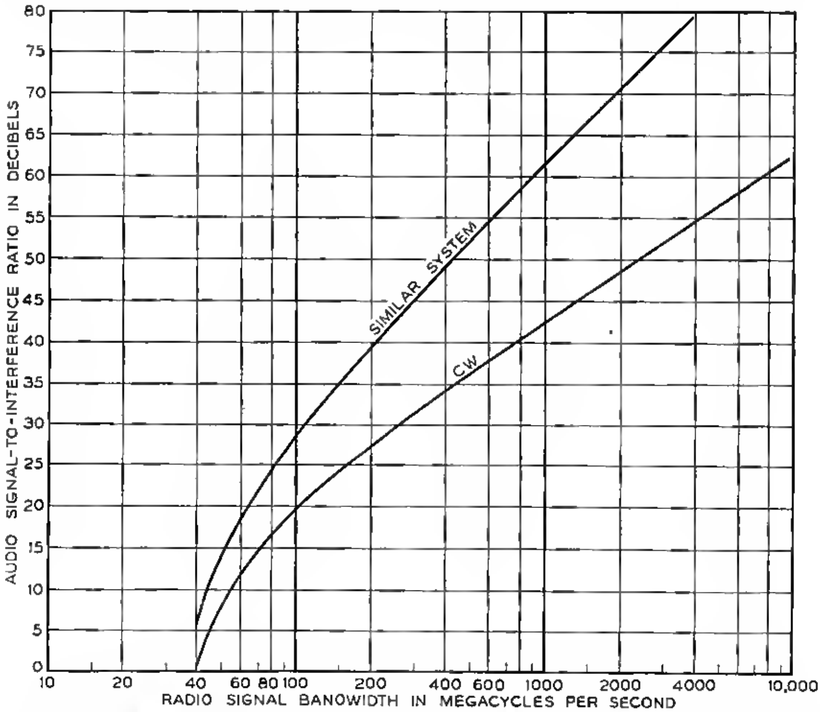


Fig. 10—PPM-AM; performance with respect to CW and similar system interference for 1000 4-kc channels with ratio of pulse to interference marginal. Relations between bandwidth and audio signal-to-interference ratio.

the power is increased with bandwidth so the slicer advantage is preserved and the slope is 6 db per octave.

The sharp reduction of signal-to-noise ratio with reduced bandwidth appearing at the left end of the curves arises from immobilizing the pulse position as the widened pulse uses up more of the total channel time allotment. According to the definition of bandwidth used here and the plan of Fig. 8, no modulation is possible when the bandwidth is $2/T$ and T is half of the channel time. For 1000 channels the channel time is 0.125 microseconds and $T = 0.0625$, which makes the audio signal-to-noise ratio zero at 32 mc.

Fig. 10—PPM-AM, CW and Similar System Interference

The curve of Fig. 10 for marginal ratio of pulse power to CW interfering power has the same shape as the corresponding curve of Fig. 9 for marginal power over fluctuation noise. There is a shift of 9 db in ordinates, however, because the peak factor of the CW interference is 9 db less than that of fluctuation noise. The interference from similar systems follows a different law because of the "exposure factor" arising from the finite probability that the interfering pulse does not overlap the wanted pulse. A straightforward probability calculation taking into account the distribution of pulse voltages in an interfering system occupying the same radio frequency band yields the curve shown on the assumption that the repetition frequencies are asynchronous. As the bandwidth is increased the pulses become shorter and their coincidences less frequent, leading asymptotically to a 9 db per octave slope instead of the 6 db per octave of the CW interference.

Fig. 11—PPM-FM, Fluctuation Noise

In the transmission of PPM by FM there are two sources of advantage over noise. One is the ordinary FM advantage and the other is the slicer advantage of PPM acting on the noise remaining in the FM output. There are, likewise, two separate conditions for system failure; one a breaking of the limiter and the other a breaking of the slicer. A certain amount of radio power will result in marginal operation of the limiter for a certain frequency swing. The corresponding deviation ratio is the quotient of the frequency swing and the baseband width; this ratio is maximum when the baseband is least. Except in the region near the minimum PPM band, advantage accrues faster with bandwidth in FM than in PPM. It is apparent, therefore, that most of the radio bandwidth should be devoted to FM advantage. The optimum proportioning occurs when the baseband width has a small value but not so small as to invoke an unsurmountable penalty by not providing for any position modulation. Mathematical analysis given in Appendix IV shows that the optimum baseband for the pulse position modulation varies with radio bandwidth in the manner shown in Fig. 11 by curve 1. Curve 2 shows the audio signal-to-noise ratio vs. radio bandwidth when the baseband width follows curve 1 and the FM limiter is marginal. It is of interest to compare curve 2 with the poorer performance of the dashed curve 3 which is calculated for the case in which both the FM limiter and the PPM slicer are marginal. The baseband width for the double marginal condition follows curve 4. Curves 5 and 6 show audio signal-to-noise ratio vs. bandwidth for constant radio power and optimum baseband. The curve of marginal amount of radio power is not given in Fig. 11, but is the same as the one given later in Fig. 13.

FIG. 12—PPM-FM, CW AND SIMILAR SYSTEM INTERFERENCE

The curve showing interference from a similar system of lower power was based on a calculation of the beat spectrum between two FM waves, both frequency modulated over the same *r-f* range (β mc) by 8 mc. The phase difference between the 8-mc modulating frequencies was assumed to vary, giving rise to various beat spectra. The power in those beat components accepted by a band zero to F_b was averaged over all 8-mc phase differences

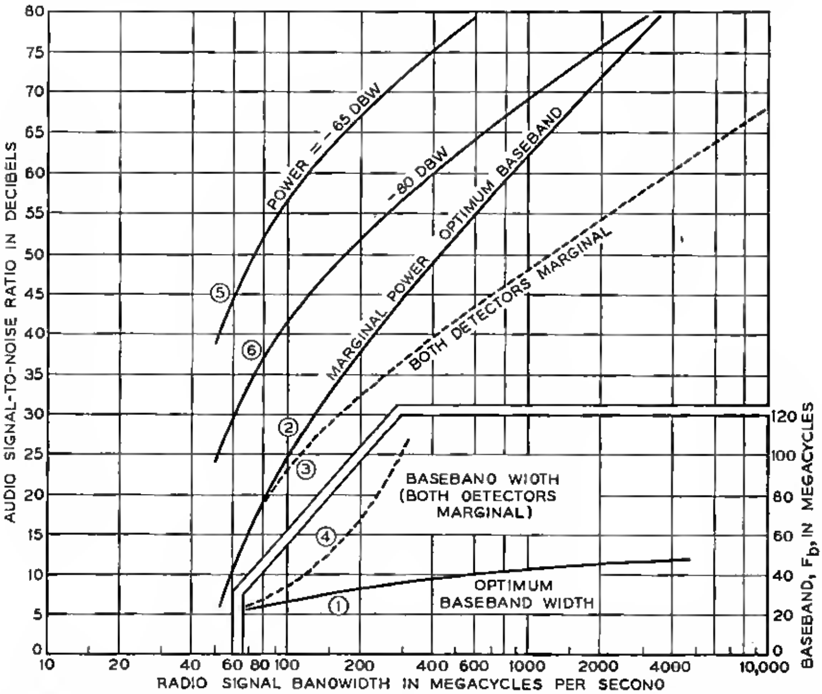


Fig. 11—PPM-FM; performance with respect to fluctuation noise. Relations between bandwidth, power, and audio signal-to-noise ratio for 1000 4-kc channels.

and this average power was taken as a measure of the interference to which the baseband signal is subjected. As outlined here, this procedure is valid for interference between idle PAM-FM systems in which the FM waves are frequency modulated as assumed above. For the PPM-FM case in which we are here interested, we take the position suggested by the transient viewpoint that the effect of interference from spaced pulses will not be much different because of their spacing and so we apply the slicer advantage possessed by the wanted system to the total interference calculated above and obtain the curve shown. At the left hand end where the pulse spacing is only slight

and the sinusoidal frequency modulation is nearly the correct representation, the above procedure is not subject to much suspicion. The validity of the right-hand portion of the curve is upheld by the fact that it is about 12 db lower than the marginal fluctuation noise curve of Fig. 11. If the wide swing FM interfering wave had a spectrum much like fluctuation noise of the same power as the FM wave the difference would be 9 db.

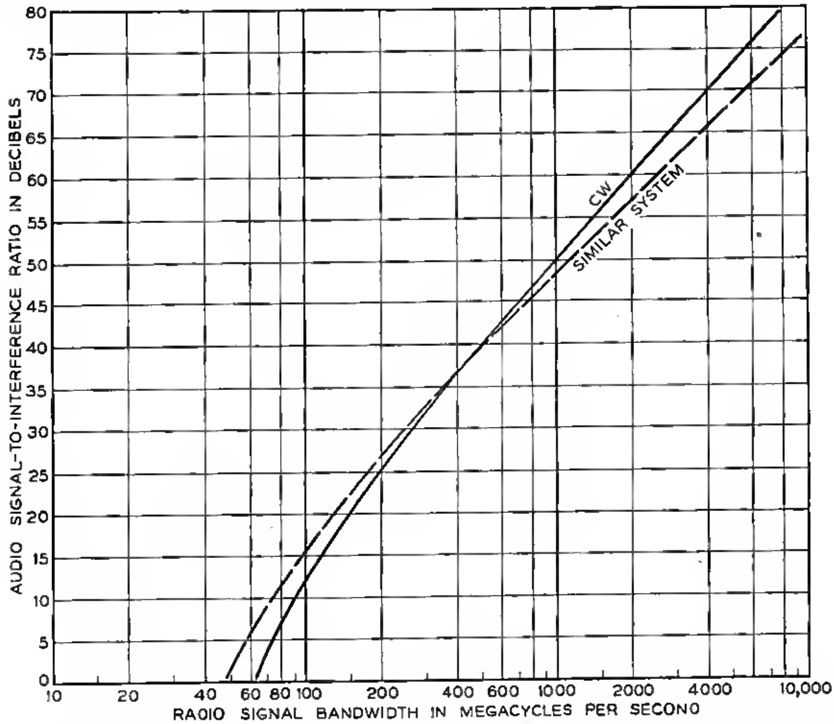


Fig. 12—PPM-FM; performance with respect to CW and similar system interference for 1000 4-kc channels with ratio of FM wave to interference marginal. Relations between bandwidth and audio signal-to-interference ratio when baseband is optimum for suppression of fluctuation noise.

It has been explicitly assumed that both systems are idle, but we see no reason to believe that if either or both were normally active the interference would be significantly different for our purposes.

Audible interference from a CW wave is caused by a disturbance to the frequency of the FM wave. Let us first assume that the CW frequency lies near the middle of the frequency swing range. No disturbance to the FM wave occurs as its frequency passes through coincidence with that of the CW but, as the frequencies diverge, the magnitude of the disturbance as well as

the frequency of the disturbance increases linearly. The baseband filter is excited only during the time the difference is less than F_b. Thus, the disturbance results from a series of perturbations to the otherwise smooth frequency variation of the FM wave. The time during which these perturbations can affect the baseband filter is short compared with the shortest pulse the baseband filter can pass, except when the baseband width is greater than half of the swing. This occurs at the extreme left-hand end of the curve. We have not attempted to calculate the response to these transients except to note that the response is a pulse which extends roughly $2T$ from its point of origin, peaking somewhere near the center of this interval. If we assume that the PPM pulses are closely spaced ($\epsilon = 0$) so that they result in a wave frequency modulated by 8 mc, there are two such evenly spaced disturbance pulses per cycle of modulation (two per $2T$ interval) and therefore there is an almost continuous disturbance wave in the base band filter output whose amplitude does not greatly exceed its RMS value. We have accordingly calculated the power sum of all the extraneous frequencies passed by the baseband filter, assuming the FM wave to be sinusoidally modulated. The location of the CW frequency giving greatest interference power was used in these calculations except in the wide band cases where the worst frequency appeared to be near the edge of the band. Here the transient viewpoint indicated that the resulting interference in the baseband would be greater if the CW frequency were nearer the center.

If the trailing edge is used to measure the time of the pulse, the principal disturbance of this time arises from the perturbation produced at the leading edge of the same pulse, and so the calculation for close-spaced pulses is not greatly in error when applied to wider-spaced pulses. If the leading edge were used the worst CW frequency for widely spaced pulses would be one differing from the rest frequency by F_b and the interference would be worse, we think, than that arising from the frequency worst for trailing edge operation.

It has been explicitly assumed that the system is idle, but we see no reason to believe that the interference would be significantly different with normal activity.

Fig. 13—PAM-FM, Fluctuation Noise

Fluctuation noise in a PAM-FM system produces the sloped noise spectrum characteristic of FM in the output of the frequency detector. The noise power per cycle is zero at zero frequency and increases with the square of the frequency. The baseband filter accepts only the portion of the spectrum between zero and F_b. If instantaneous sampling of the signal values is used, all noise frequencies in this range are equally effective as causes of errors. Use of a channel gate of maximum permissible duration

consistent with a satisfactory margin over crosstalk from adjacent channels furnishes a practical method of discriminating against the influence of noise components near the top of the baseband where the noise spectrum is strongest. The exact shape of the gate is not very critical. The curves have been calculated for a rectangular gate coincident with the channel allotment time, which is just possible without crosstalk in the case of non-

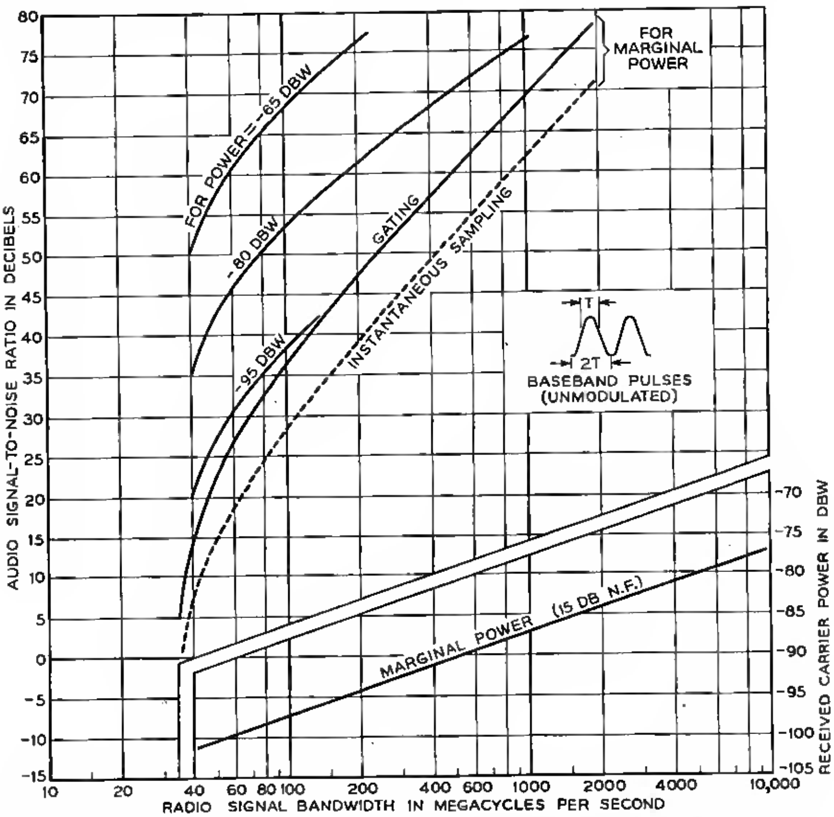


Fig. 13—PAM-FM; performance with respect to fluctuation noise. Relations between bandwidth, power, and audio signal-to-noise ratio for 1000 4-kc channels.

overlapping sinusoidal pulses. A somewhat shorter rectangular gate or a gate of sinusoidal shape leads to very nearly the same results. The advantage of gating as compared to instantaneous sampling is approximately 8 db. Calculation of the gated noise is a straightforward process if based on the concept of the FM noise spectrum acting as signal on a product demodulator in which the carrier consists of the harmonics of the gating function. Each harmonic demodulates the spectrum centered about the harmonic

frequency and contributes audio power proportional to tbe product of harmonic power and spectral density. The channel filter accepts only the demodulated noise falling in the audio signal range.

The marginal power curve has been drawn for a 3 db ratio of peak carrier to peak interference or 12 db ratio of mean carrier power to mean fluctuation noise power. Curves for specific amounts of received power are included as

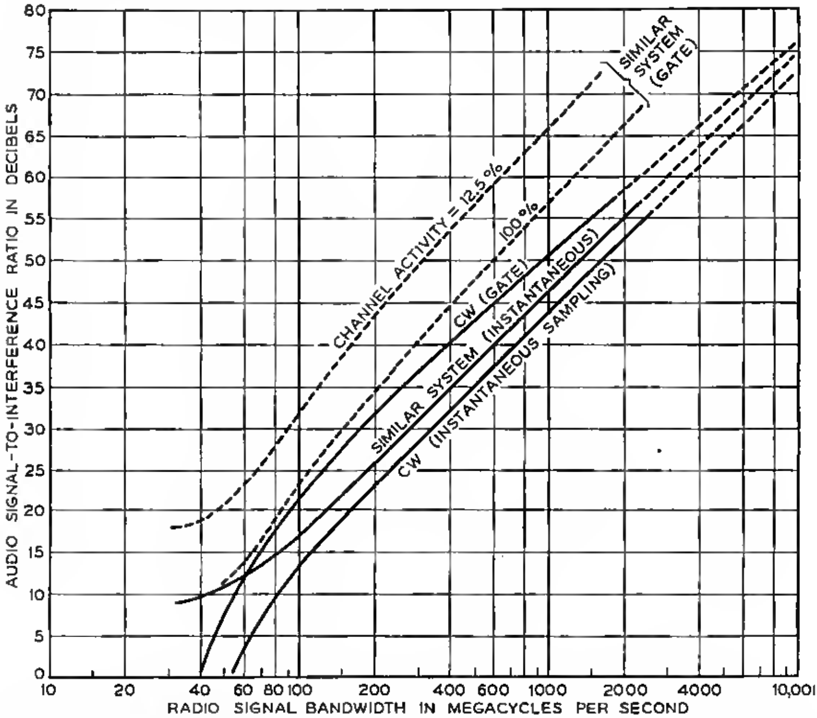


Fig. 14—PAM-FM; performance with respect to CW and similar system interference for 1000 4 kc channels with ratio of FM wave to interference marginal. Relations between bandwidth and audio signal-to-interference ratio.

well as the curve of marginal received power vs. radio signal bandwidth for a receiver with 15 db noise figure.

Fig. 14—PAM-FM, CW and Similar System Interference

CW interference can be calculated conveniently by assuming all channels idle and determining at what frequency within the radio signal band a CW component of fixed power produces maximum disturbance of an audio channel. This worst possible amount of disturbance is then assumed not to be much affected by the various channel loading conditions existing during normal operation of the system. When all channels are idle, the transmitted

carrier of a PAM-FM system using sinusoidal pulses assumes the particularly simple form of an FM wave modulated by a sinusoidal signal having frequency Nf_r (8 mc for 1000 channels) and total frequency swing $\beta/2$. Hence the rigorous steady state solution for interference between CW and sinusoidally modulated FM was calculated and the interfering components falling in the baseband range selected. The gating function was then applied to these components in the same way as described above for fluctuation noise, and the resulting products falling in the audio channel range evaluated. The signal-to-interference ratio was expressed as the ratio of rms signal power received from a full-load channel test tone to the rms value of the audio interference. A range of frequency locations for the CW interference was investigated for each radio signal bandwidth and the one giving maximum audio interference used for the point on the curve. The worst position of the CW was usually found to be near the extremities of the idle channel frequency swing. Curves are shown for a rectangular gate of maximum duration and for instantaneous sampling.

When the source of interference is a similar system, we assume that the midband frequencies differ only slightly. With both systems idle, we have two sinusoidally modulated FM waves which are identical except for (1) a small variable difference between mean carrier frequencies and (2) a variable phase shift between the two modulating frequencies. The interference falling in the baseband consists of steady state components which are approximately harmonics of the channel slot frequency Nf_r. As is characteristic of FM interference, the amplitude at the mth harmonic contains a factor proportional to m; and the component near zero frequency, the approximate zeroth harmonic, is very small. If we gate this interference with a rectangular gate of duration $1/Nf_r$, we find that the gated output vanishes for input components at Nf_r, $2Nf_r$, ... , because these frequencies are located at the infinite loss points of the aperture admittance. The gate would transmit the zeroth harmonic, but this component tends toward zero amplitude. Our conclusion is that two idle PAM-FM systems accurately lined up to occupy the same frequency range are balanced against interference from each other when a rectangular channel gate of full channel allotment time is used. The balance tends to disappear as the channels are loaded because the interference then spreads throughout the base band instead of being concentrated at the blind spots of the aperture. Thus, for the first time in our consideration of pulsed systems, we are obliged to take account of channel loading conditions.[23]

[23] A wave could be frequency modulated about a central frequency by PAM pulses of plus and minus sign and an idle system would thus consist of a wave of constant frequency. The weaker of two such idle systems aligned in frequency would produce no (or very little) interference in the other, using either channel gating or instantaneous sampling. The susceptibility to CW interference would be greater than in the biased modulation assumed above, however.

We note that the balance disappears if instantaneous sampling is used instead of gating because there is no longer any aperture discrimination. The curve for instantaneous sampling is plotted on Fig. 14. Calculation of this curve brings out the fact that the amplitude of the interfering components also depend on the framing phase difference between the two systems. When the framing frequencies are in phase, the two waves have a constant frequency difference, and the interference vanishes. We assume the phase difference as equally likely to fall anywhere within a complete cycle and average the received interference power over all phases. The curve is found to approach an asymptotic ordinate of 9 db at minimum band width as the frequency swings on the two systems approach zero together. The 9 db limit is compounded of 3 db from the marginal ratio between the two carriers, 3 db from averaging over the carrier phase difference, and 3 db from averaging over the framing phase difference. When wide bands and large swings are used, the curve approaches parallelism with the dashed one of Fig. 13 for fluctuation noise but about 15 db lower. Of this difference 9 db is accounted for by the higher marginal mean power level. The remainder is assignable to differences in spectral distribution; in particular the *r-f* spectrum of idle similar system interference is concentrated in half of the band instead of being uniformly spread as in fluctuation noise.

The curve for gated similar system interference has been estimated by assuming that, with all channels loaded and a wide frequency swing, the performance is like that with fluctuation noise except for a 3 db correction allowed for the more concentrated spectrum. This gives an asymptote on the right parallel to the solid marginal curve of Fig. 13 and 12 db lower. At the left the curve must approach the same asymptote as the one for instantaneous sampling. It then seems reasonable to assume that the interference power is directly proportional to the number of active channels and the curve for an average of one eighth of the channels loaded is obtained by raising the full load curve 9 db.

Fig. 15—PCM-AM

The curves on the left show how the audio signal-to-noise ratio varies with bandwidth, the audio noise being quantizing, or granularity, noise as discussed in Appendix I. The number of digits per code symbol is n and the number of digit values (including zero), i.e., the base, is b. Bandwidth is $2/T$ where T is the time per pulse and is therefore 16 mc per digit. The curves are, of course, a set of discrete points rather than continuous as shown. The steep rise is to be contrasted with the 3 to 9 db per octave slopes of the curves previously presented.

The curves on the right plot the maximum values (with a 3 db allowance included) of peak noise or interference, referred to the highest pulse value,

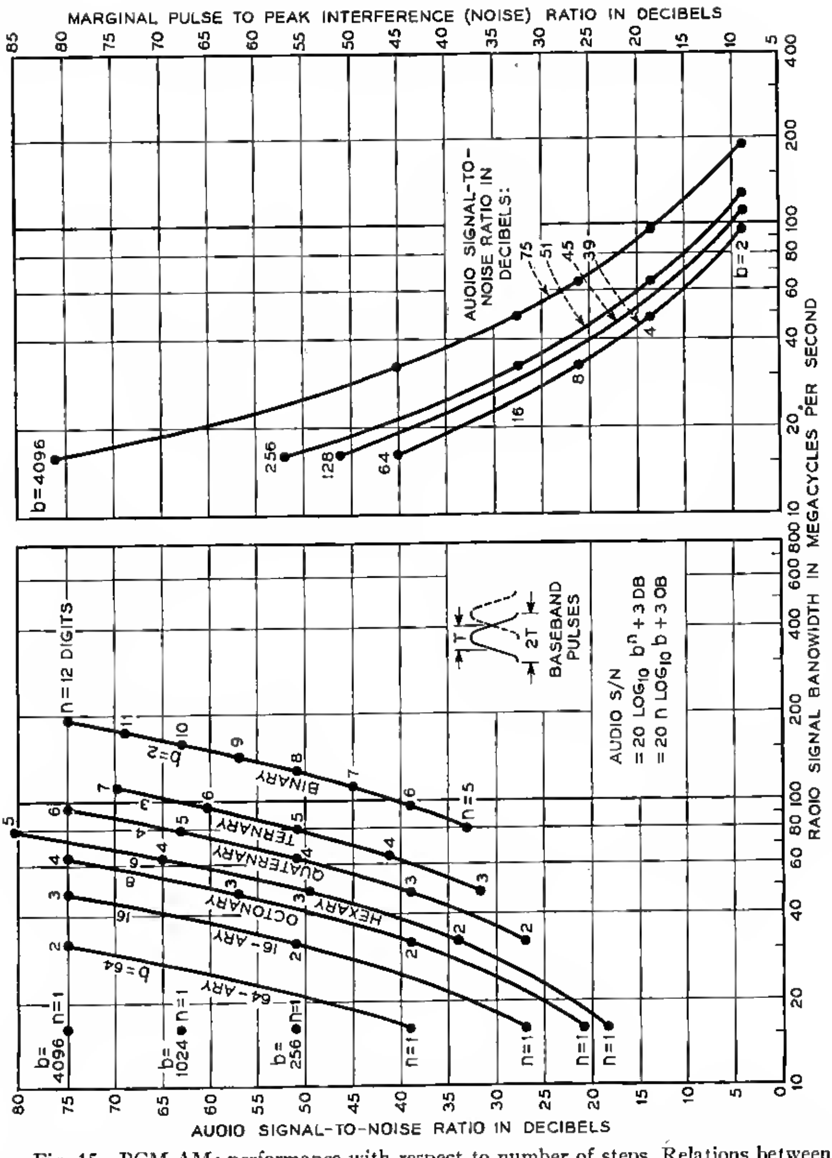


Fig. 15—PCM-AM; performance with respect to number of steps. Relations between bandwidth and audio signal-to-noise ratio and between bandwidth and required pulse-to-interference ratio for 1000 4-kc channels.

permitted by the system, versus bandwidth for three different audio signal-to-noise ratios. The lower boundary of the curves corresponds to binary PCM while the left-hand boundary corresponds to the other extreme, namely

quantized PAM having the number of steps necessary to yield the specified signal-to-noise ratio. The quantized PAM bandwidth of 16 mc assumes the use of overlapping sinusoidal pulses as in binary PCM. Actually, such an overlap would be hazardous in the higher base systems; and quantized PAM, like unquantized PAM, should perhaps be assigned more time per pulse but not as much as $2T$ because regeneration could be employed to prevent accumulation of interchannel crosstalk. The tables presented later do not include the bandwidth increase that would follow such an increase in time per channel.

The curves at the right in Fig. 15 are terminated at 16 mc corresponding to one pulse per channel. In accordance with the principles of Appendix III more than one channel per pulse can be transmitted, theoretically. To include such a hypothetical case of less than one digit per channel, the curves could have been extended upward to the left. The 39 db signal-to-noise ratio curve would have reached an ordinate of 81 db at 8 mc on the bandwidth axis.

It is of interest to compare the audio signal-to-noise ratio of unquantized PAM with that of quantized PAM for the interference ratios demanded by quantized PAM. In the case of marginal CW interference the audio noise (evaluating the audio disturbance as noise of equivalent power) turns out to be the same as the quantizing noise and so, in a circuit of more than one span, quantized PAM is advantageous from a transmission point of view. With fluctuation noise the unquantized PAM audio noise would be 9 db lower than the quantizing noise and so, in a circuit of more than 9 spans of equal loss, the quantized PAM would be preferred.

Fig. 16—PCM-FM, Fluctuation Noise

Here FM advantage is employed to permit operation in the presence of more noise than is possible with AM. It seems more illuminating to explain these curves by checking their correctness rather than by deriving them.

In all cases, a baseband signal-to-noise ratio giving the same margin over noise peaks as for AM (Fig. 15) is obtained by FM advantage. For the solid curves the FM limiter is assumed to be marginal (12 db radio signal-to-noise ratio), and for the dashed curves the radio signal-to-noise ratio is assumed to be the same as the marginal requirement for binary PCM-AM (18 db). The FM advantage with respect to an FM wave of the same power as in the peak AM pulse is, in db

$$20 \log \frac{\beta}{F_b} + 4.8 = 20 \log \left(\frac{B}{F_b} - 2\right) + 4.8$$

However, the FM power is greater than the peak AM pulse power by 10

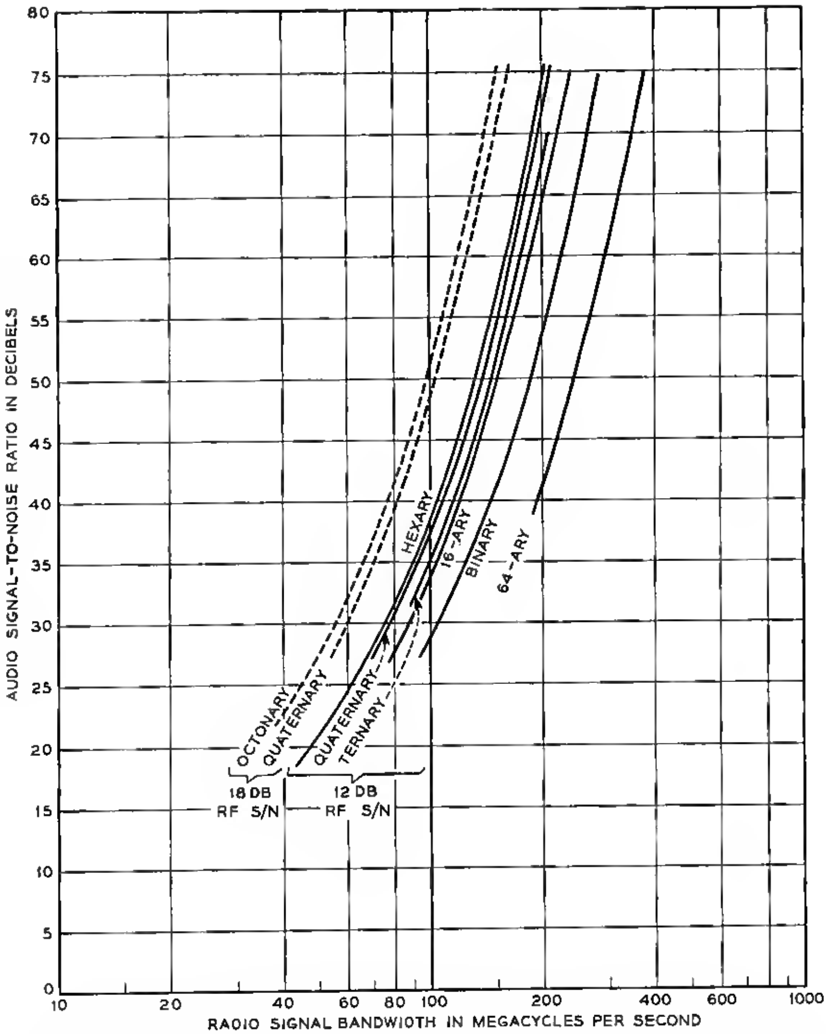


Fig. 16—PCM-FM; performance with respect to fluctuation noise. Relations between bandwidth, power, and audio signal-to-noise for 1000 4-kc channels, showing optimum bases for different ratios of FM wave to noise.

$\log \frac{B}{2F_b}$ because the FM power-to-noise ratio is maintained regardless of bandwidth. The total advantage of the FM case is therefore

$$20 \log \left(\frac{B}{F_b} - 2 \right) + 4.8 + 10 \log \frac{B}{2F_b}$$

and this must just make up for the difference between the 12 db (or 18 db) FM wave-to-noise ratio and the pulse-to-noise ratios of 18 db + 20 log $(b - 1)$ required in the AM case. Substituting values from the curves will show that this is so.

These curves show a minimum bandwidth for an optimum PCM base. This is to be expected since two different rates of excbange between bandwidth and advantage are involved. One is the advantage growing out of PCM of reduced base while the other is the conventional FM advantage. An analogous situation was found in PPM-FM.

It is of interest to examine the PCM-FM situation when the FM circuit is as tolerant of noise as the most tolerant AM case, namely when tbe *r-f* signal-to-noise ratio is 18 db. The optimum PCM base is octonary and the corresponding minimum bandwidth (as we define it) is actually 20% less than for binary AM. This apparent advantage of PCM-FM is not obtained when tolerance to CW and similar systems is considered. Figure 17, which follows, shows that when allowance is made for a 9 db *r-f* signal-to-interference ratio (as in binary PCM-AM), the minimum FM bandwidth is greater by about 30% than for binary AM and the optimum base is ternary or quaternary. If the 3 db interference tolerance possible in FM is required, it is obtained, as shown in Fig. 17, witb ternary PCM-FM, at a cost of approximately twice tbe bandwidth required in binary PCM-AM, which bas a tolerance of 9 db. We should point out here that binary PCM transmitted by single sideband and detected by a local carrier has a tolerance of 3 db and requires half the bandwidth sbown in Fig. 16. PCM-FM requires a bandwidth 3.8 times that of single sideband binary PCM for tbe same 3 db tolerance.

FIG. 17—PCM-FM, CW AND SIMILAR SYSTEM INTERFERENCE

In PCM, sequences of several pulses of the same amplitude may occur. The FM signal then consists of a steady frequency. A steady beat frequency persisting for several pulse periods will be produced by CW interference.[24] If this beat frequency is F_b the maximum interfering amplitude will be produced. The amplitude is $(Q/P)\ F_b$ while tbe step interval is $\beta/(b - 1)$. To confine the interference to a half step (witb 3 db margin) requires tbat

$$\beta/(b - 1) \geqq 2(Q/P)\sqrt{2}\, F_b$$

For $Q/P = 0.707$,

$$\beta \geqq 2\,(b - 1)\, F_b$$

[24] The general solution of the problem of frequency error produced by superimposing a sine wave on an unmodulated carrier is given in Appendix II.

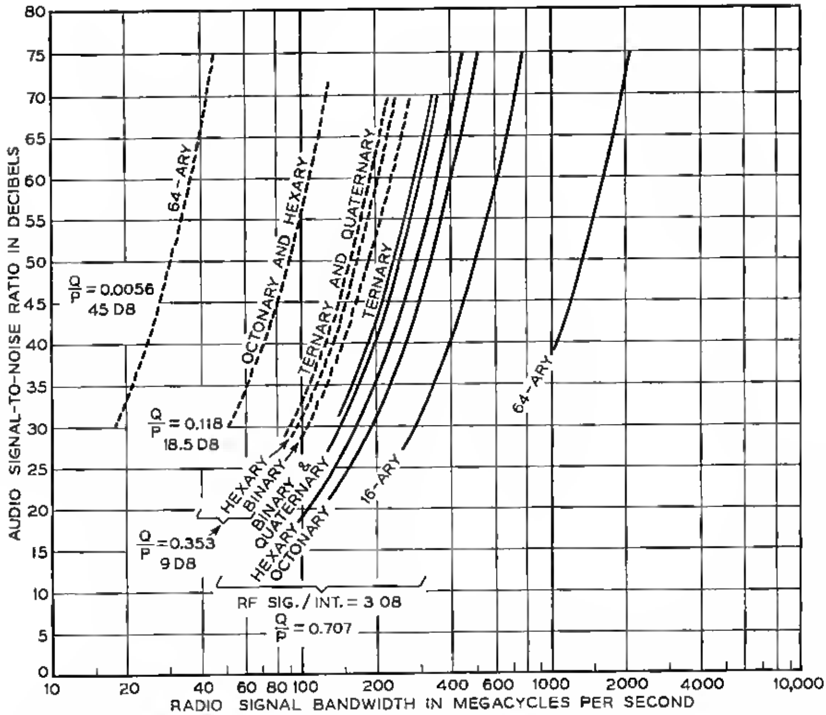


Fig. 17—PCM FM; performance with respect to CW and similar system interference for 1000 4-kc channels. Relations between bandwidth and audio signal-to-noise ratio showing optimum bases for different ratios of FM wave to interference.

and the minimum band width becomes

$$B = \beta + 2\,F_b = 2bF_b$$

For $Q/P = 0.3535$ (9 db) the minimum required value of β is halved so that

$$B = (b + 1)\,F_b$$

For $Q/P = 0.118$ (18.5 db) the minimum required value of β is further reduced threefold so that

$$B = \frac{b + 5}{3}\,F_b$$

The curves of Fig. 17 are calculated from these relations.

With interference from a similar system a number of necessary conditions must be met. If the systems are similar in PCM base and in radio frequency, the sustained beat frequencies that may occur are $\beta/(b - 1)$ and

multiples thereof. The amplitude of the lowest of these frequencies is $(Q/P)\beta/(b-1)$. For $Q/P = 0.707$, this frequency must be suppressed by the baseband filter since otherwise the threshold would be exceeded. Thus,

$$\beta \geqq (b-1)\,F_b$$
$$B = \beta + 2F_b = (b+1)\,F_b$$

For $Q/P = 0.353$, the lowest beat frequency need not be suppressed but the $2\beta/(b-1)$ frequency must be suppressed; thus $2\beta/(b-1) \geqq F_b$.

$$\beta \geqq \frac{b-1}{2}\,F_b$$

$$B = \beta + 2F_b = \frac{b+3}{2}\,F_b$$

Comparing these bandwidth values with those required for CW shows that the above requirements are more lenient than for the corresponding CW cases, particularly for the higher values of b where the above bandwidth values approach one-half of those obtained for CW.

However, the above requirements are not quite sufficient. Transitions between adjacent frequency values, occurring in one system, will produce varying beat frequencies which pass through all values. This case differs from the CW case in that the beat frequency is not sustained and that the baseband filter output will not be as high as in the CW case. Calculations show that the bandwidth requirements are intermediate between those for the CW case and the similar system case considered previously.

When we remember that for low base systems (binary) the requirements for similar system and CW interference are nearly alike, while for high base systems a small frequency difference in frequency alignment can produce similar system interference completely equivalent to CW interference, we may regard Fig. 17 as applying to both, practically. Such a conclusion also makes the curves apply to interference between systems of different base.

Quantized PPM

PCM pulses, including the limiting case of quantized PAM pulses, may be transmitted by time modulation instead of frequency modulation, i.e., by "quantized PPM." In this case, as in PCM-FM, bandwidth may be used to increase the tolerance to noise and interference. Figure 18 illustrates this case. At (A) is shown a PCM pulse having b values including zero, the highest amplitude being unity. The maximum tolerable peak interference is $1/2\sqrt{2}\,(b-1)$ and the time per pulse is taken as T. (For high base systems the time per pulse should be greater, perhaps $2T$ as pointed out in the discussion of PCM-AM). At (B) is shown the quantized PPM

pulse of the same amplitude. The time per pulse is taken to be T'' and the duration of the pulse at half height is T'. If T'' is put equal to T the ratio of the PPM bandwidth to the PCM bandwidth is T/T'. The time shift

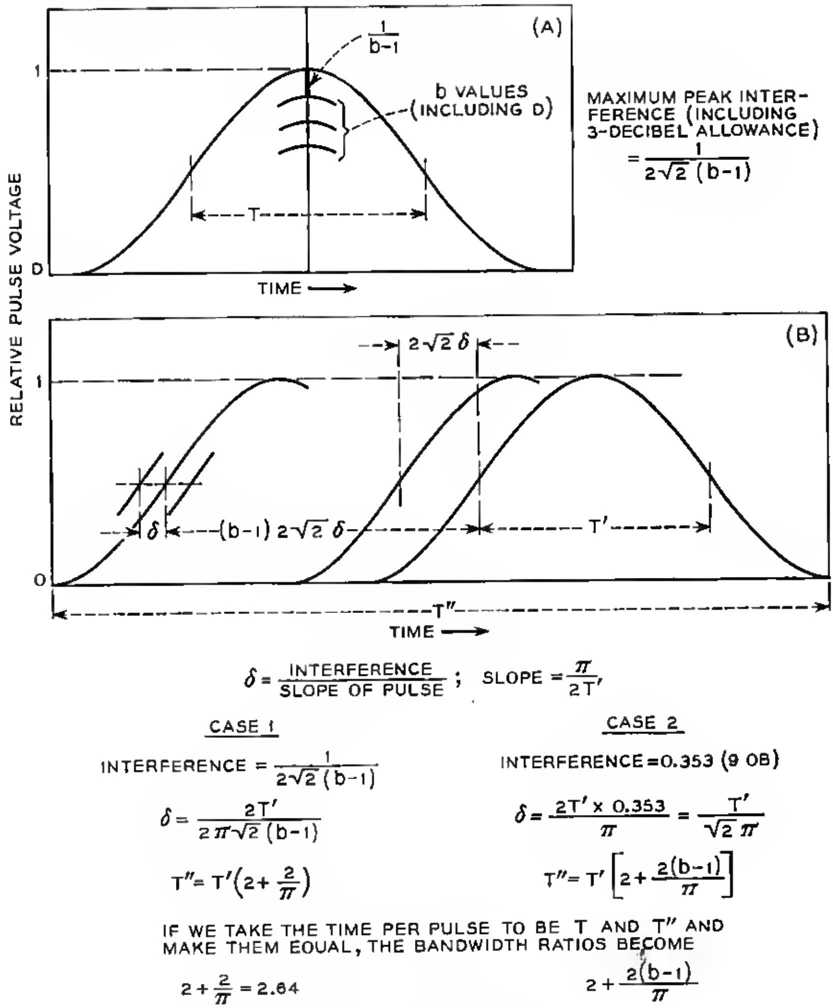


Fig. 18—Comparison of quantized PAM with quantized PPM.

produced by peak interference is represented by δ. Our system is marginal when the smallest quantized step produces a shift of $2\sqrt{2}\,\delta$. The peak time shift produced by a signal is then $(b-1)2\sqrt{2}\,\delta$. Two cases are considered. In one the PPM system operates in the presence of the same peak interference as in the PCM case. The bandwidth ratio is 2.64. The other case

assumes that the peak interference is marginal for all bandwidths (9 db down). The bandwidth ratio is then $2\left(1 + \frac{b-1}{\pi}\right)$. It was previously found that in PCM-FM, with peak interference 9 db down, the radio bandwidth must be $(b + 1)F_b$. Since the radio bandwidth of PCM-AM is $2F_b$, the bandwidth ratio is $(b + 1)/2$. Comparing these bandwidth ratios we see that the PPM bandwidth required to operate in the presence of marginal interference is nearly two times that required in PCM-FM. Furthermore, this PPM bandwidth ratio applies to marginal fluctuation noise whereas in PCM-FM a more favorable result was obtained.

Fig. 19—FDM-FM, Fluctuation Noise

When a group of channels in frequency division is transmitted by frequency modulation, the addition of channel voltages is translated to an addition of instantaneous frequency shift. The non-simultaneous load advantage applicable to a multichannel amplifier for frequency divided channels thus becomes an advantage in reduction of total frequency swing as compared to the sum of the individual peak frequency swings of the channels. The numerical db increments versus number of channels listed in Table I should, however, be modified for the following reason: The fluctuation noise spectrum in the output of an FM detector is not uniform with frequency, and hence the noise is unequally distributed among the channels. In order to obtain the same noise in all channels it is necessary to taper the signal levels in such a way that the full load frequency swing produced by one channel is proportional to the frequency of the channel. The frequency swing corresponding to full load in the top channel is therefore a larger part of the maximum instantaneous swing required for the group than the swings corresponding to lower channels. The result is, in effect, phase modulation. The multiplex addition factors for tapered level channels have not been determined experimentally. We have assumed here a 3 db reduction in the power capacity values listed in Table I. These reduced values then give the incremental capacity referred to *full load on the top channel*. Curves are shown for 100, 500 and 1000 channels. On account of the multiplex addition factor, it is not possible to obtain results for other numbers of channels from one curve by simply changing the frequency scale.

The derivation of these curves is straightforward but leads to an expression for the required bandwidth as a root of a cubic equation. As in the case of Fig. 16 we shall discuss the FDM-FM curves by checking them numerically. We have assumed that the channels are tapered in level and that we have, in fact, phase modulation with its consequent flat baseband noise distribution. To check the 60-db point on the 1000-channel

curve, we calculate that the noise in the entire baseband must be 43 db below the power in a sine wave which employs the full system load capacity. This figure comes from reducing the 60 db full load *channel* ratio by 30 db because of the 1000-fold greater baseband width and increasing it by the amount, $16 - 3 = 13$ db, by which the full system load must exceed full load in the top channel. Thus $60 - 30 + 13 = 43$ db. An FM advantage of 31 db must be obtained to permit the marginal *r-f* signal-

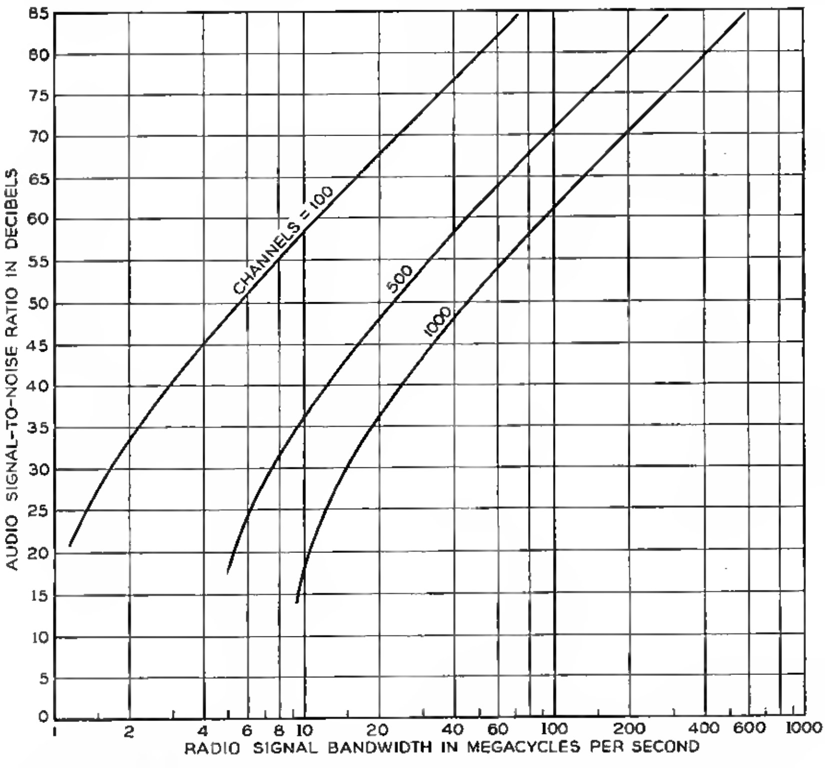


Fig. 19—FDM-FM; performance with respect to fluctuation noise. Relations between bandwidth and audio signal-to-noise ratio for marginal power; 4-kc channels.

to-noise ratio of 12 db to satisfy the above requirement, $(43 - 12 = 31)$. We get this advantage in part by phase modulation gain given by $20 \log \frac{\beta}{F_b} - 6$ db. This gain is referred to 100% modulation AM whose unmodulated carrier power is the same as the FM wave power. This means that the reference is a system in which the FDM baseband appears as upper and lower sidebands which, when demodulated, yield a baseband

signal-to-noise ratio equal to the ratio of unmodulated carrier power to the noise power in the double width radio band. Since we keep the FM power marginal for all bandwidths an additional bandwidth improvement of $10 \log \frac{B}{2F_b}$ accrues. Substituting $B = 92$ mc, $F_b = 4$ mc, and $\beta = 92 - 8 = 84$ mc, will show that the above gains total 31 db.

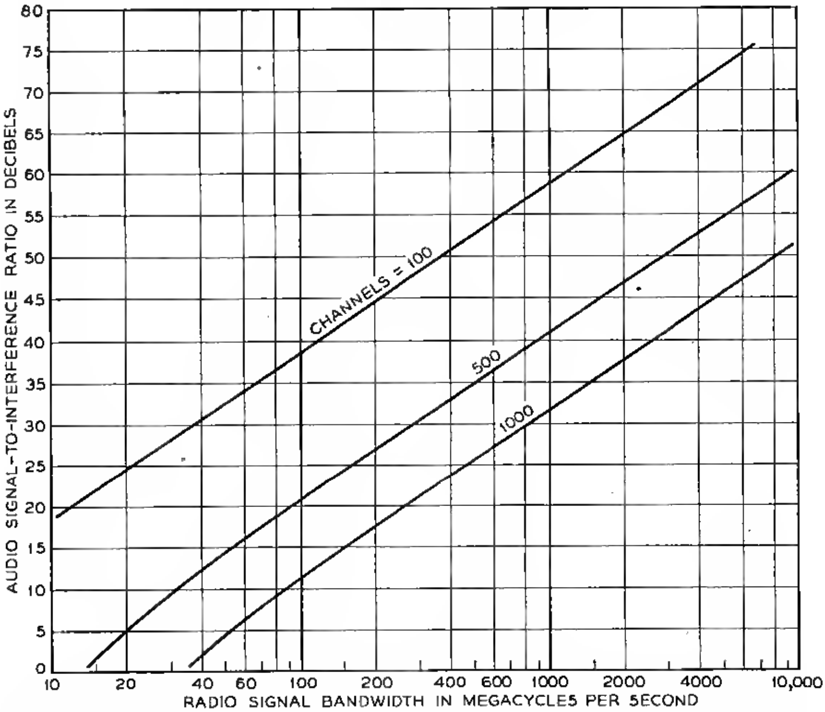


Fig. 20—FDM-FM; performance with respect to CW interference. Relations between bandwidth and audio signal-to-interference ratio for marginal ratio of FM wave to interference; 4-kc channels.

Fig. 20—FDM-FM, CW Interference

The disturbance produced by CW is most readily evaluated when all channels are idle for then we have only the frequency error produced by a sine wave of relatively small amplitude superimposed on the steady sinusoidal carrier wave. To a first approximation (see Appendix II) the error has a frequency equal to the difference between the carrier and CW frequencies and an amplitude equal to this frequency difference multiplied by the ratio of the CW to the carrier amplitude. The error thus increases linearly with the frequency of the channel in which it falls but, since the channel levels

are also tapered in the same way, the signal-to-interference ratio is independent of the frequency of the disturhing CW. Varying the CW frequency only changes the numher of the channel into which the interference falls. Loading the channels distrihutes the interference over several channels instead of concentrating it in one, hut we have plotted in Fig. 20 the more severe case in which all channels are idle.

We have not undertaken to compute curves for similar system interference in the case of FDM-FM, but estimates for two extreme conditions can he made. In the case of low index FM systems the carrier frequency component of the spectrum is not affected by the modulating signal and the FM wave is, in fact, like an AM wave with the carrier displaced 90 degrees in phase. A similar interfering FM wave comhines with the wanted FM to produce frequency or amplitude variations and does this cyclically as the *r-f* phase between the systems varies. When the phases are appropriate for the production of frequency variation, crosstalk appears in the wanted reception at a level lower hy the ratio of FM wave amplitude. Averaging over all *r-f* phases should reduce the crosstalk hy 3 dh. The actual amount of interference received in a channel is less than would he predicted from replacement of the interfering FM wave by fluctuation noise of the same mean power spread over the *r-f* band, because the hulk of the interfering power is contained in the carrier component located at a frequency which does no harm. Increase of the frequency swing in both systems produces significant reduction in crosstalk when the carrier amplitude diminishes appreciably and important higher order sidebands appear, i.e. when the interfering system has its spectrum spread out more or less uniformly, like noise. Systems designed for wide swings under full load may, however, operate with only a few channels active; in such cases the low index situation may exist and the received interference will he down approximately by the ratio of the FM waves, without the henefit of FM advantage. While in this situation the hulk of the interfering power is again contained in the harmless carrier, the received interference is concentrated in a few channels and is greater than if the interfering wave power were spread, like noise, thinly over the *r-f* band, which in this case is many times wider than the band occupied by the low index signal. For such adverse loading conditions, the curve for similar system interference, while starting at the left ahove the corresponding noise curve of Fig. 19, may actually cross over and finally approach it from the lower side.

In the case of systems of very wide swing such as are involved in Table II we regard the interfering system as equivalent, under all common load conditions, to noise spread uniformly over the bandwidth and having the same power as the interfering wave. The entry in Table II is obtained hy reading

the curve of Fig. 19 at 69 db audio signal-to-noise ratio which would be appropriate to yield 60 db when the "noise" is marginal at 3 db below the FM wave power instead of 12 db. A different procedure is required for the narrow band entries of Table IV. Here tbe emphasis on conservation of bandwidth leads to a two-frequency repeater plan with tolerance of similar system interference 44 db down. A 60 db audio signal-to-interference ratio can be met under these conditions with moderate swings for which the equivalent noise representation of the interference is not valid. The result is considerably influenced by the channel loading and we have no impeccable method of calculating the necessary bandwidth. We estimate that a bandwidth of 22.5 mc., with $\beta = 14.5$ mc., will satisfy the requirements for all except unusually adverse loading conditions.

III. Band Width and Power Tables

The information contained in the curves of Fig. 9–20 has been used in preparing Tables II and III, which show what can be done with the various systems when bandwidth is used freely. The prime objective studied here is the conservation of peak transmitted power. In Table II the audio channel must meet message circuit requirements[25] while, in Table III, a much better grade of performance—more than sufficient for transmission of high fidelity musical programs—is stipulated. We have prepared Table III (as well as Table V) on the basis of replacing the 1000 4-kc. message channels of Table II with 250 16-kc. channels. Since we have available established load rating theory only for message circuits, we bave omitted FDM and FDM-FM from Table III (and Table V). The values of Table II are based on a nominal 60 db ratio of signal-to-noise, but it is assumed permissible to meet this in the pulse systems by using 22 db of instantaneous companding so that only 38 db signal-to-noise ratio is actually required within the compandor. The PCM systems provide for a 39 db circuit within the compandor, corresponding to 6 binary digits, 3 quaternary digits, 2 octonary digits or one 64-ary digit. No allowance is made for the accumulation of quantizing noise arising when several PCM links are connected in tandem at voice frequency. In practice, 7 binary digits might be used. This would provide for several links and would permit slightly more companding. Table III assumes a 75 db signal-to-noise ratio and no companding and is referred to here as a "program" circuit. We use such a high-grade circuit to illustrate more emphatically how the system preferences depend

[25] We do not pretend to deal fully with the involved matter of system requirements distinguishing between kinds of noise and interference or crosstalk that appear in message channels. We merely assume that the power of the separate types of disturbances considered must be individually 60 db below that of a full load test tone under the worst specified transmission condition.

TABLE II

OPTIMUM BAND WIDTHS FOR MINIMUM POWER FOR MESSAGE TYPE CIRCUITS

133 30-mi spans, 1000 4-kc channels, 15 db NF, 75 db span loss.

SYSTEM	S/N(1) IN DB	NUMBER OF RESHAPINGS: (4) 133 — BAND WIDTH IN MC	133 — POWER IN WATTS	5(2) — BAND WIDTH IN MC	5(2) — POWER IN WATTS	1(3) — BAND WIDTH IN MC	1(3) — POWER IN WATTS	ONE EXPOSURE TO MARGINAL INTERFERENCE, BAND WIDTH IN MC — CW	ONE EXPOSURE TO MARGINAL INTERFERENCE, BAND WIDTH IN MC — SIMILAR SYSTEM	MARGINAL PEAK INTERFERENCE IN DB
PPM-AM	38	2000 / 1000 / 0	10 / 5 / 0; 0.6	2000 / 1000 / 0	10 / 5 / 0	2000 / 1000 / 0	10 / 5 / 0	2000 / 1000 / 0	2000 / 1000 / 0	-9
PPM-FM	38	2000 / 1000 / 0	1.0 / 0.5 / 0; 0.05	2000 / 1000 / 0	1.0 / 0.5 / 0	2000 / 1000 / 0	1.0 / 0.5 / 0	2000 / 1000 / 0	2000 / 1000 / 0	-3
PAM-FM	38	500 / 250 / 0	1.0 / 0.5 / 0; 0.03	500 / 250 / 0	1.0 / 0.5 / 0	500 / 250 / 0	1.0 / 0.5 / 0	500 / 250 / 0	500 / 250 / 0; (5)	-3
FOM-FM	60	500 / 250 / 0	1.0 / 0.5 / 0; 0.03	500 / 250 / 0	1.0 / 0.5 / 0	500 / 250 / 0	1.0 / 0.5 / 0	500 / 250 / 0; 25,000	500 / 250 / 0	-3

(1) AUDIO SIGNAL TO NOISE (OR INTERFERENCE) RATIO
(2) RESHAPING EVERY 27 SPANS
(4) REGENERATION IN THE CASE OF PCM. LIMITING IN THE CASE OF FM

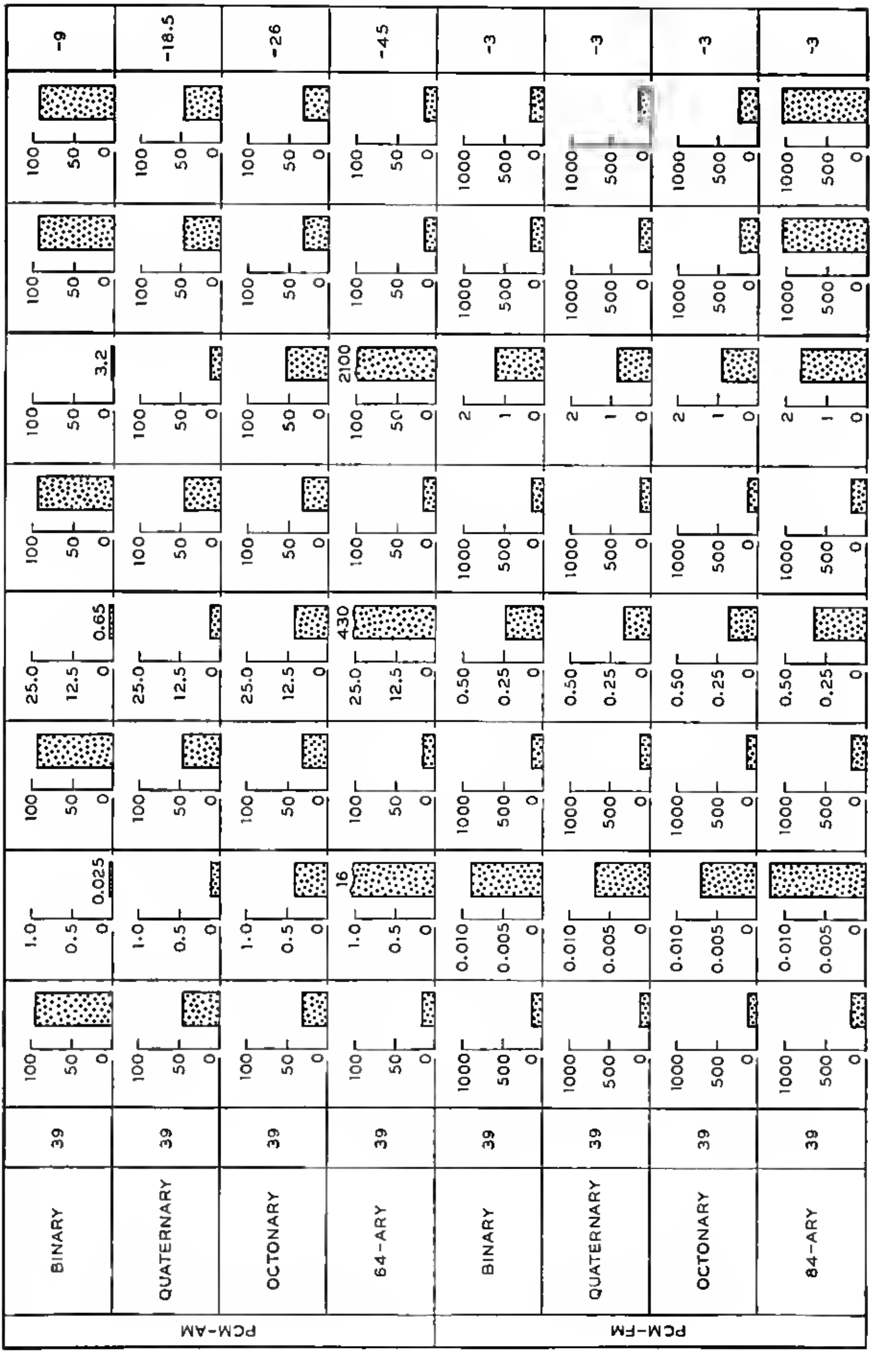

PCM-AM
PCM-FM
BINARY
QUATERNARY
OCTONARY
64-ARY
84-ARY
39
-9
-18.5
-26
-45
-3
0.025
16
0.65
430
3.2
2100

TABLE III

OPTIMUM BAND WIDTHS FOR MINIMUM POWER FOR PROGRAM TYPE CIRCUITS

133 30-mi spans, 250 16-kc channels, 15 db NF, 75 db span loss.

SYSTEM	S/N(1) IN DB	NUMBER OF RESHAPINGS:(4) 133		5(2)		1(3)		ONE EXPOSURE TO MARGINAL INTERFERENCE BAND WIDTH IN MC		MARGINAL PEAK INTERFERENCE IN DB
		BAND WIDTH IN MC	POWER IN WATTS	BAND WIDTH IN MC	POWER IN WATTS	BAND WIDTH IN MC	POWER IN WATTS	CW	SIMILAR SYSTEM	
PPM-AM	75	100,000 50,000 0	500 250 0 43	100,000 50,000 0	500 250 0	100,000 50,000 0	500 250 0	100,000 50,000 0	100,000 50,000 0 2800	-9
PPM-FM	75	10,000 5000 0	20 10 0 0.8	10,000 5000 0	20 10 0	10,000 5000 0	20 10 0	10,000 5000 0	10,000 5000 0	-3
PAM-FM	75	10,000 5000 0	20 10 0 0.5	10,000 5000 0	20 10 0	10,000 5000 0	20 10 0	10,000 5000 0	10,000 5000 0 (5)	-3

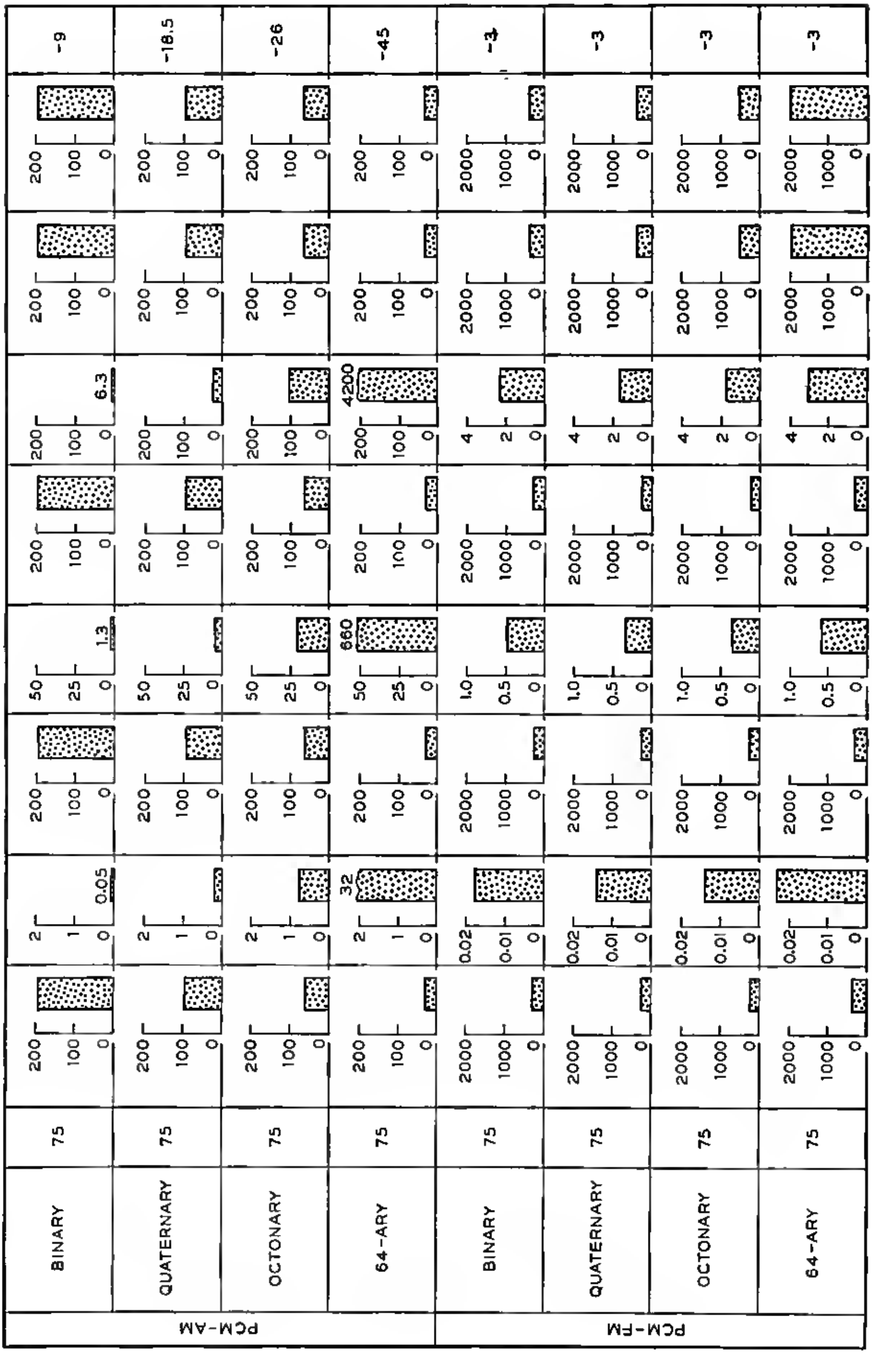
(1) AUDIO SIGNAL TO NOISE (OR INTERFERENCE) RATIO

(2) RESHAPING EVERY 27 SPANS

(3) [illegible]

(4) REGENERATION IN THE CASE OF PCM, LIMITING IN THE CASE OF FM

(5) [illegible]


PCM-AM
BINARY
QUATERNARY
OCTONARY
64-ARY
PCM-FM
BINARY
QUATERNARY
OCTONARY
64-ARY
75
0.05
32
1.3
660
6.3
4200
-9
-18.5
-26
-45
-3
-3
-3
-3

upon signal-to-noise ratio. Both tables apply to long systems comprising 133 repeaters spaced 30 miles apart. The span loss is 75 db. Three reshaping or regeneration plans are shown: reshaping at every repeater, reshaping five times in the complete system, and no reshaping within the system.

In the case of PCM, reshaping becomes a true regeneration which completely removes noise accumulated in the transmission link; while, in the non-quantized pulse systems, reshaping restores the original pulse shape but retains timing or amplitude errors. In the case of FM systems, reshaping is accomplished by the amplitude limiter which removes envelope variations arising from noise, but does not suppress the accompanying frequency shifts. Reshaping, in contrast to regeneration, is only a partial prevention of cumulative effects but, as shown by the tables, it has definite value in enabling the use of wider transmission bands with corresponding smaller amounts of power than permissible without reshaping. Regeneration completely removes errors in both the amplitude and in the time. The maximum bandwidth which can be used is independent of the number of regenerations, but the signal power must be increased in proportion to the number of spans covered before regeneration.

In the case of non-regenerative radio transmission systems, we may regard the 75 db span loss as 60 db free space loss plus a fading allowance of 15 db. This allowance is intended to cover the increment of noise caused by fading in the entire system. Available data on distribution of fading are too meager to permit generalization, but indicate that on some routes, at least, the total degradation suffered would rarely be worse than that produced by 15 db simultaneous fades on all spans; and hence a design based on 75 db span loss should satisfy the noise requirements except for an extremely small fraction of the time. In other words a few spans of the non-regenerative system may fade deeply but, in regard to total accumulated noise, the system is credited with the higher signal-to-noise ratio occurring on spans which are not simultaneously in fading minima. In regenerative PCM systems no credit accrues from a higher signal-to-noise ratio occurring between points of regeneration, and protection must be provided against the worst condition that is likely to occur in any section included between regeneration points.

However, the values of minimum power obtained by lavish use of bandwidth exhibited in Tables II and III may be particularly significant for wave guide transmission systems; and hence the assumption that all spans have the same loss is appropriate here.

The outstanding features of Tables II and III are the extremely small amounts of power needed in the PCM systems with only moderate expenditures of bandwidth as compared with the non-quantized system. These

results illustrate the properties of PCM as a means of converting bandwidth into transmission advantage.

The PCM-FM entries are taken from the curves of Fig. 16 plotted for noise 12 db down. Curves are also given in Fig. 16 for noise 18 db down. It will be noted that the bandwidths indicated become smaller when the noise is required to be farther down, but that the power requirements become greater because the bandwidth reduction factor is less than the factor multiplying the *r-f* signal-to-noise ratio.

The columns at the right in Tables II and III show the bandwidtbs which must be employed in order to attain a 60 db signal-to-interference ratio in the presence of one source of interference whose amplitude is just marginal for the type of system concerned.

Tables IV and V are prepared from another point of view—that of conserving bandwidth[26] instead of power. The systems particularly suited for narrow bands such as FDM and PAM-AM have been added to the list. The actual minimum bandwidths are, in many cases, determined by engineering judgment; smaller values than those tabulated may be possible at the expense of greatly increased power and precision requirements. Thus in the case of PPM-AM we have arbitrarily chosen 40 mc as necessary for 1000 4-kc channels. According to our initial postulate the audio signal-to-noise ratio vanishes at 32 mc, and indefinitely great signal power would be required as we approach this limit. In PAM-AM we have assumed that pulses in adjacent channels just touch, thereby setting the bandwidth at 32 mc. Smaller bandwidths could be used if the pulses were allowed to overlap. This would reduce the allowable duration of the channel gate and deprive the system of some of its tolerance to similar system interference as well as noise. The maximum pulse power required for 100% modulation is tabulated. If instantaneous sampling were used this would be 6 db above the unmodulated pulse power which is, in turn, 38 db above the mean total fluctuation power accumulated in a 32-mc band from 133 spans. We have reduced the value of power thus computed by 1.7 db to allow for a calculated improvement in signal-to-noise ratio obtainable by gating at the channel input with a time function of the same shape as the signal pulse.

The FM systems listed are of two kinds. The first is a relatively narrow-band type in which advantages such as relative immunity to gain fluctuation and amplitude non-linearity are sought with small increment in bandwidth over AM. Since these objectives are not sufficient in themselves to fix the actual bandwidth needed, an arbitrary additional requirement has

[26] We do not here entertain the idea of using certain exchange methods to permit use of less band width than the conventional minimum of 4 kc per channel, but rather to use modest amounts of additional band width. Appendix III discusses briefly a band reduction principle.

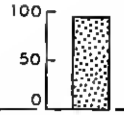
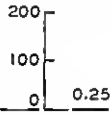
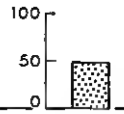
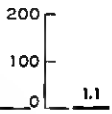
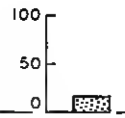
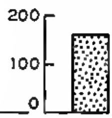
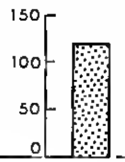
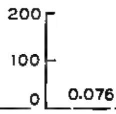
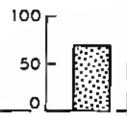
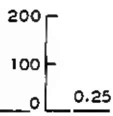
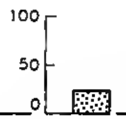
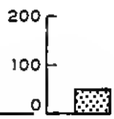
TABLE IV

MINIMUM BAND WIDTHS AND CORRESPONDING POWER REQUIREMENTS FOR MESSAGE TYPE CIRCUITS

133 30-mi spans, 1000 4-kc channels, 15 db NF.

SYSTEM		S/N IN DB	BAND WIDTH IN MC	POWER IN WATTS	TOLERABLE INTERFERENCE RATIO IN DB		SPAN LOSS IN DB
					CW	SIMILAR SYSTEM	
PPM-AM		38		850	46	41 (1)	75
PPM-FM		38		850	49	39 (1)	75
PAM-AM		38			44	40 (1)	75
PAM-FM	NARROW BAND	38			46	23 (2)	75
PAM-FM	WIDE BAND	38		1.2	20	9 (2)	75
FDM		60			76	60	75
FDM-FM		60			68	44	75

TABLE IV—*Concluded*

SYSTEM		S/N IN DB	BAND WIDTH IN MC	POWER IN WATTS	TOLERABLE INTERFERENCE RATIO IN DB		SPAN LOSS IN DB
					CW	SIMILAR SYSTEM	
PCM-AM	BINARY	39			9	9	85
	QUATERNARY	39			18.5	18.5	85
	64-ARY	39			45	45	85
PCM-FM	QUATERNARY	39			9	9	85
	OCTONARY	39			18.5	18.5	85
	64-ARY	39			45	45	85

REGENERATE AT EVERY REPEATER IN PCM

(1) INTERFERING SYSTEM IDLE; ACTIVITY HAS SMALL EFFECT.
(2) ASSUMING 12.5% CHANNEL ACTIVITY.

Table V

Minimum Band Widths and Corresponding Power Requirements for Program Type Circuits

133 30-mi spans, 250 16-kc channels, 15 db NF.

SYSTEM		S/N IN DB	BAND WIDTH IN MC	POWER IN WATTS	TOLERABLE INTERFERENCE RATIO IN DB		SPAN LOSS IN DB
					CW	SIMILAR SYSTEM	
PPM-AM		75		4.3×10^6	83	78(1)	75
PPM-FM		75		4.3×10^6	86	76(1)	75
PAM-AM		75		1.4×10^6	81	77(1)	75
PAM-FM	NARROW BAND	75		1.4×10^6	83	60(2)	75
PAM-FM	WIDE BAND	75	1270	18.8	25	9(2)	75
PCM-AM	BINARY	75		0.5	9	9	85
PCM-AM	QUATERNARY	75			18.5	18.5	85
PCM-FM	QUATERNARY	75		0.15	9	9	85
PCM-FM	OCTONARY	75			18.5	18.5	85

REGENERATE AT EVERY REPEATER IN PCM

(1) INTERFERING SYSTEM IDLE; ACTIVITY HAS SMALL EFFECT.

(2) ASSUMING 12.5% CHANNEL ACTIVITY.

been imposed that the transmitted power in the FM system should be equal to that required in the corresponding AM system. The resulting values of bandwidth are found to be reasonable ones for low index FM. An exception is made in the case of FDM-FM systems of message type where it is found that a net economy of frequency occupancy can be obtained by increasing the frequency swing sufficiently to tolerate similar system interference 44 db down. This enables the two-frequency repeater plan of Fig. 4, as discussed in Section I, to be used and substantially reduces the frequency occupancy over that of a lower index FM system more vulnerable to antenna crosstalk and therefore requiring more frequency assignments. We have estimated that a radio signal bandwidth of 22.5 mc achieves the required 44 db tolerance. The second type of FM is a wide-band system designed for specified tolerances of interference from similar systems. Data on this second type will be used later in our study of inter-route interference, where it will appear that ruggedness is a more important criterion of frequency occupancy than the minimum bandwidth needed for transmission.

No curves have been furnished to determine the FDM entries, since there is no variation with radio bandwidth to consider. The band required is merely the number of channels multiplied by the width of a channel. The power required for message channels is determined by calculating the amount of power in one channel to give a 60 db margin over mean fluctuation noise power in a 4-kc band and applying the multiplex addition factor of Table I. Similar system interference is simply linear crosstalk and must be, we say, 60 db down. CW interference referred to maximum system power must be down an additional amount equal to the multiplex addition factor of Table I in order to meet 60 db suppression in the disturbed channel. Since the two-frequency plan of Fig. 4 does not suppress interference between the two directions of a single route by 60 db, we must use twice as many frequency assignments as there shown. This duplication will appear in Table VI. FDM is the only system of Table IV for which such duplication is necessary, since the others do not require more than 44 db suppression. In the program type systems of Table V, however, the first four listed would need duplicated frequency assignments.

The PCM-AM systems of course do not use any smaller bandwidths than those given in Tables II and III and would, therefore, be expected to show disadvantageously in a bandwidth comparison with the other systems. On the other hand they make, relatively, a good showing in power requirements and in tolerance to CW and similar system interference.

In the next section we shall show that economy of bandwidth may, in fact, be illusory because of the greater susceptibility to intra- and inter-system interference associated with narrow band methods. It is not the bandwidth actually needed for transmission that is important, but the

tightness with which the bands may be packed into the frequency range without mutual interference. Because of the ruggedness of low-base PCM, neighboring frequency bands can actually be allowed to overlap. Introduction of the proper spacing factors for satisfactory separation in frequency between adjacent bands causes the PCM system to overtake the other methods in effective utilization of frequency space, especially when intersecting routes are involved.

The PCM-FM systems listed in Tables IV and V are of the second class listed above, in which equivalent ruggedness rather than equivalent power as compared to AM is the criterion. Thus, binary PCM-AM is compared with a PCM-FM system having the same 9-db tolerance of interference. The curves of Fig. 17 show that, with such a tolerance, the minimum PCM-FM bandwidth is secured when the base is either three or four. We choose the quaternary case here because the signal-to-noise ratios obtainable coincide with those of the binary. Likewise, either octonary or hexary PCM-FM furnishes the optimum base for the 18.5 db tolerance possessed by quaternary PCM-AM and, of the two, octonary is more suitable for our tabulation. In determining the power required to override fluctuation noise in a PCM-FM system designed for a specified tolerance of similar system interference, we must make sure that both the limiter and slicer are protected against breaking.

The values of repeater power capacity shown in Tables IV and V will satisfy the noise requirements on a 133-span non-regenerated circuit with 75 db loss on all spans. For spans of 60 db free space loss the tabulated power thus provides for 15 db fades simultaneously on all spans, or for 13 db simultaneous fades of 25 db, or for a single fade of 36 db. PCM systems employing regeneration on every span must be powered for the deepest fade that is likely to be encountered. We have arbitrarily taken this to be 25 db making the span loss 85 db. This is probably not a sufficient allowance for some situations but will serve for illustrative purposes. If regeneration were not practiced the power would be $25 - 15 = 10$ db lower from the fading allowance standpoint but would have to be increased $10 \log 133 = 21$ db for noise accumulation, so regeneration results in a power saving of 11 db. If, with regeneration, we were to protect each span against the deepest single fade (36 db) permitted by the power provided for non-regenerative operation, the power advantage of regeneration would disappear ($36 - 15 - 21 = 0$ db).

In general, when the power without regeneration protects against simultaneous fades upwards of just a few db there is little or no power advantage in regeneration if we then protect each span against the deepest single fade permitted when regeneration is not practiced. This is true even for large numbers of spans. There remain, however, important advantages for

regeneration in preventing accumulation of disturbances that are not much affected by the distribution of fading.

IV. Frequency Occupancy Tables For Radio Relay

The frequency space occupancy for a single two-way route is, according to principles laid down in the introduction, a frequency block $2U$ times the signal bandwidth.[27] Our problem, as stated in the introduction, is to examine the situations arising when a number of 1000-channel routes converge toward a terminal city, assuming all of the routes to be of the same kind. We will determine the number of times the above frequency blocks must be repeated in the spectrum in order to keep interference within tolerable bounds. The sum of these blocks then really defines the frequency occupancy and determines the space which must be allocated or, conversely, determines the number of routes a given allocation will accommodate. We will use the tolerable ratios of similar system interference taken from Tables IV and V, together with appropriate antenna directivity, to determine the number of these blocks.

Antenna Characteristics

The directional discrimination afforded by the antennas is obviously an important factor in frequency economy. For our present study, we employ an antenna having a directional pattern slightly superior to that of the 4000-mc shielded lens antenna in use on the New York-Boston radio relay circuit. Figure 21 shows the assumed directional characteristic omitting "nulls" between the minor lobes. Of importance also are the nearby discrimination characteristics of the antennas as given in Fig. 4.

The situations arising at a point where a number of routes converge (or cross) or where a route is equipped with a spur connection are variations of that occurring at a single repeater point. In fact, the situation in which two routes converge from approximately opposite directions occurs at every repeater point in a straight route, while a repeater point at which the route bends sharply is like a terminal point at which two routes converge at a small angle.

The crosstalk in our assumed two-frequency long distance repeater system has been estimated in Section I (under "The Radio Repeater") and was found equivalent to a single source of similar system interference 44 db down. A system which possesses just enough tolerance to withstand the accumulated crosstalk on a long straight repeater system is not capable of meeting another such system at an angle unless additional frequencies are

[27] In the case of very tender systems, such as FDM, the factor 2U is replaced by 4U because a four-frequency plan is needed for a two-way repeater. U is the band spacing factor discussed under "The Radio Repeater."

invoked. The FDM-FM system of "minimum band width" listed in Table IV is intended to possess this 44 db tolerance. To illustrate the requirement of new frequencies let us take the case of several FDM-FM routes

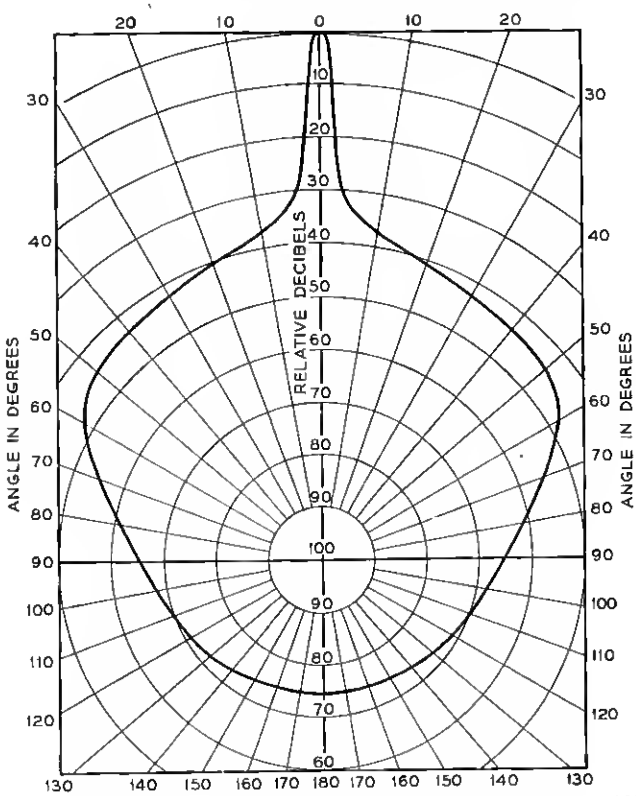


Fig. 21—Assumed directional selectivity of 10 ft. x 10 ft. antenna at 4000 mc.

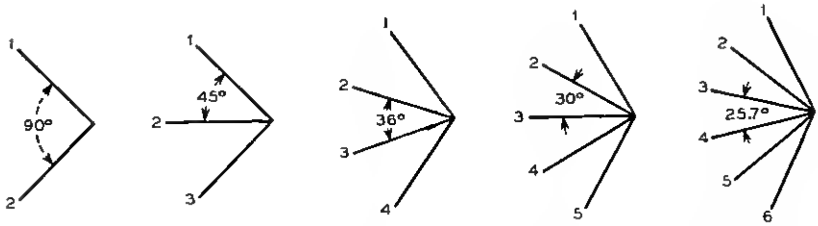


Fig. 22—Simplified route pattern for study of selectivity required in congested localities.

converging in the manner suggested by any of the diagrams of Fig. 22. A critical situation occurs in regard to the frequencies used for receiving at the point toward which the systems converge. If the same receiving

frequency were to be used on two or more routes, receiving directional discrimination amounting to 75 db would have to be secured:

1. Required interference ratio	44 db from Table IV FDM-FM	
2. Allowance for repeater crosstalk	1	(51 db down)
3. Differential fading allowance[28]	30	
	—	
	75 db	

The repeater crosstalk is here taken to be equivalent to one source 51 db down which is the value corresponding to no differential fading on adjacent spans as calculated in Section I. It will be remembered that allowance was made for a single differential fade of 30 db occurring somewhere along the route. Here we assume that this differential fade may occur between two of the converging paths and we demand that the receiving directional discrimination shall protect the system against such an occurrence. In this case the required directional discrimination turns out to be equal to the 75 db front-back ratio from which the 44 db figure was obtained. This is manifestly impossible with the assumed directivity characteristic[29] and the angles involved. Therefore, different receiving frequencies are required on each route. These same frequencies may be used for transmitting at the junction, provided the disposition of terminals is such as to provide enough directional discrimination and physical separation to permit operation at the low received level in the face of the high transmitted level on the same frequency. The interference path loss plus antenna discrimination must be, for the case involving the longest span:

1. Required interference ratio	44 db (FDM-FM, Table IV)
2. Allowance for repeater crosstalk	1
3. Free space span loss	60
4. Fading allowance	25
	—
	130 db

We continue our discussion of the converging routes of Fig. 22 by assuming that:

1. Conditions encountered elsewhere on the routes do not restrict the freedom to switch the frequencies among the routes.

2. The disposition of terminals at the junction is such that inter-terminal interference is not a controlling factor.

Under the above assumptions the directional discrimination of the terminal

[28] This differential fading allowance corresponds to a fade of 25 db below free space on one route and a 5 db increase over free space on the other.

[29] The use of perpendicular polarizations cannot, we assume, be counted on to give further discrimination when the directional discrimination is already 40 or more decibels.

antennas alone determines the number of different frequencies. The same receiving frequency may be used at the junction on two routes separated by an angle sufficient to yield the required antenna discrimination. All intervening routes must employ different frequencies. The frequencies so determined may be used for transmitting from the junction if they are staggered with respect to the receiving frequencies. Take, for instance, the five-route plan shown in Fig. 22. Suppose the directional discrimination needs to be 60 db for a particular system. The directional pattern shows that this requirement is met at 85 degrees. Thus, routes 1 and 4 may use frequency A, say. Routes 2 and 3 then must use different frequencies, B and C. Thus, we have:

Route	1	2	3	4	5
Trans. Freq.	A	B	C	A	B
Rec. Freq.	B	C	A	B	C
or	(C	A	B	C	A)

While the treatment of the route congestion problem outlined above is oversimplified it enables us to make a broad survey having some significance.

Table VI for 1000 4-kc message channels and Table VII for 250 16-kc "program" channels were derived on the above basis. The decibel figures at the head of each column are the allowable interference ratios from Tables IV and V increased by 30 db for differential fading.[30] A single source of interference of the values given in the table is supposed to degrade the circuit to the minimum requirements for a long circuit. In regenerative PCM there is no accumulation of degradation due to interference occurring on various spans. In non-quantized systems such degradations are cumulative. However, when protection to the above values is provided, with no allowance of 30 db for differential fading of the desired and interfering signals, the occurrence of *simultaneous* additional degradations is extremely unlikely. Protection against this severe fading at one point alternately protects against the simultaneous occurrence of several less severe fades.

The values of repeater power capacity shown in the table will satisfy the noise requirements on a 133-span transcontinental nonregenerated circuit with 15 db fades simultaneously on all spans. This is equivalent to providing for 13 simultaneous fades of 25 db or, statistically, for the fading that is not likely to be exceeded except during a small fraction of the time. The PCM systems employing regeneration are powered for 25 db fades on any or all spans.[31] The free space span loss is 60 db.

In computing the frequency occupancy we take cognizance of the fact

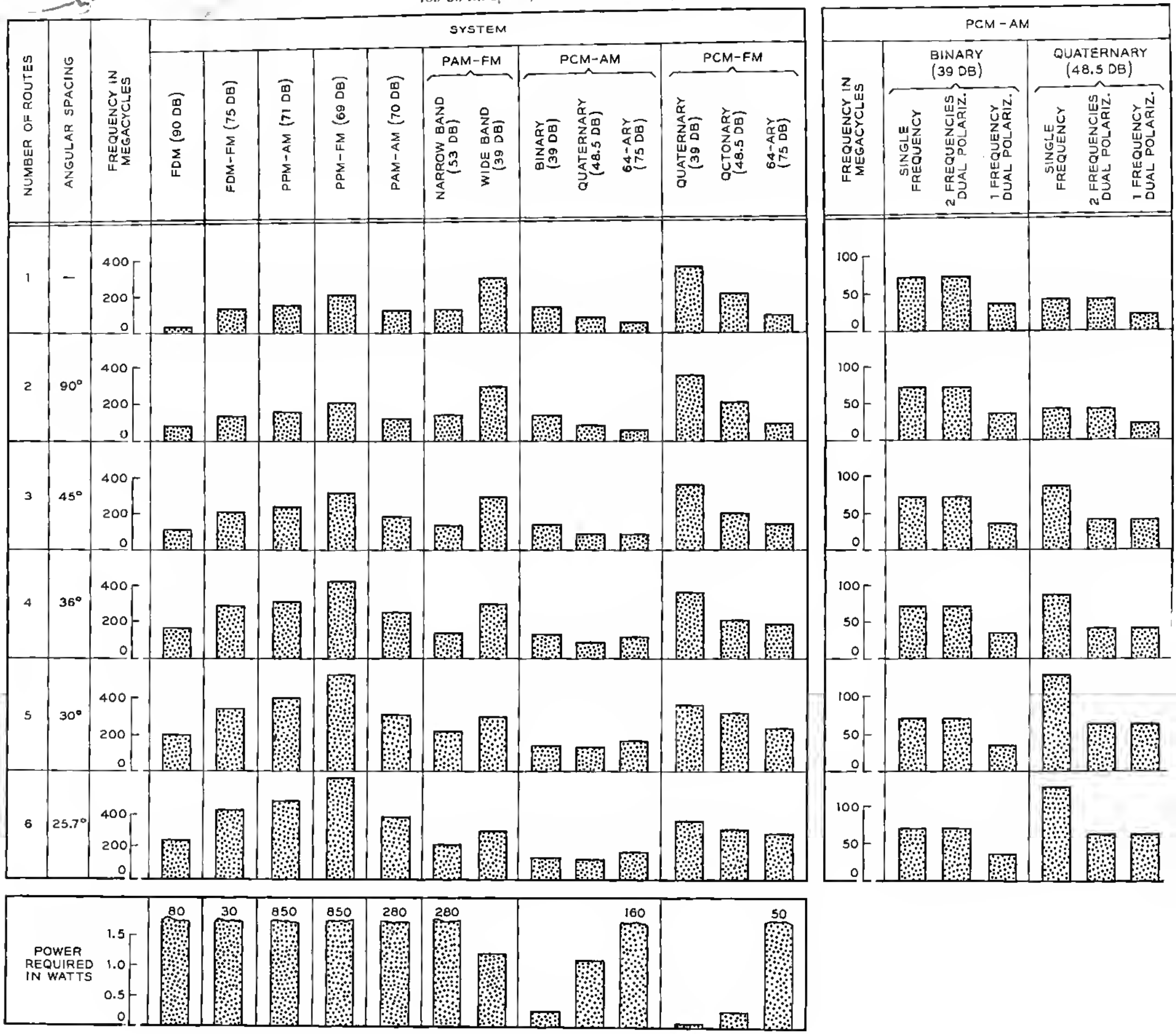
[30] The FDM-FM entry provides 1 db additional allowance for repeater crosstalk as mentioned before. Repeater crosstalk is negligible in the other systems.

[31] No such distinction was made in Tables II and III. There, provision was made for 75 db span loss in all cases.

TABLE VI

TRUE FREQUENCY OCCUPANCY OF VARIOUS MESSAGE GRADE RADIO RELAY SYSTEMS FOR CONGESTED ROUTES

133 30-mi spans, 1000 4-kc channels, 15 db NF.



that systems of different vulnerahility require different guard space to protect against adjacent-hand crosstalk. In Section VII it is concluded that with binary PCM-AM the band spacing could perhaps be as small as $1.5/T_0$ with realizahle filters and gates. T_0 is the time allotted to one pulse. Underlying the meaning of B as used heretofore in connection with PCM-AM, is the relation $T_0 = 2/B$. If the band spacing is to he taken as $1.5/T_0$, the hand spacing may he expressed in terms of B as 0.75 B. In other words, the band spacing factor U may be reduced to 0.75.

In the case of ideal *FM* systems the receiving frequency discrimination need not suppress adjacent radio signal hands to anywhere near tbe degree required of co-channel interference, provided the near edges of the adjacent signal bands differ hy more than the widtb of the basehand filter. We do not, of course, assume ideal apparatus and have heen rather liheral in guard space allowance.

The following band spacing factors, to he used with the handwidths of Tahles IV and V, are considered realizable and consistent with tbe shape of the spectrum to be transmitted. A reduction to practice would likely lead to somewhat different factors hut tbese will suffice for our illustrative Tahles VI and VII. In the non-regenerative systems, tbe spacing factor required to protect against interference from new frequencies required at a junction is, perhaps, less than is required to protect against interference at every repeater on a long route. A small economy in occupancy could properly he invoked on this account in some cases hut, in tbe interest of simplicity, this has been neglected, and the same factor U will he associated with every frequency required.

System of Table VI	Factor U
FDM	2.5
FDM-FM	3
PPM-AM	2
PPM-FM	2
PAM-AM	2
PAM-FM (narrow band)	2
PAM-FM (wide band)	1.5
PCM-AM (64-ary)	2
PCM-AM (quaternary)	0.9
PCM-AM (binary)[12]	0.75
PCM-FM (64-ary)	2
PCM-FM (octonary)	1.5
PCM-FM (quaternary)	1.5

[12] The experimental system described by Meacham and Peterson (loc. cit.) employs a spacing factor of 1.12.

Table VII

True Frequency Occupancy of Various Program Type Radio Relay Systems for Congested Routes

133 30-mi spans, 250 16-kc channels, 15 db NF.

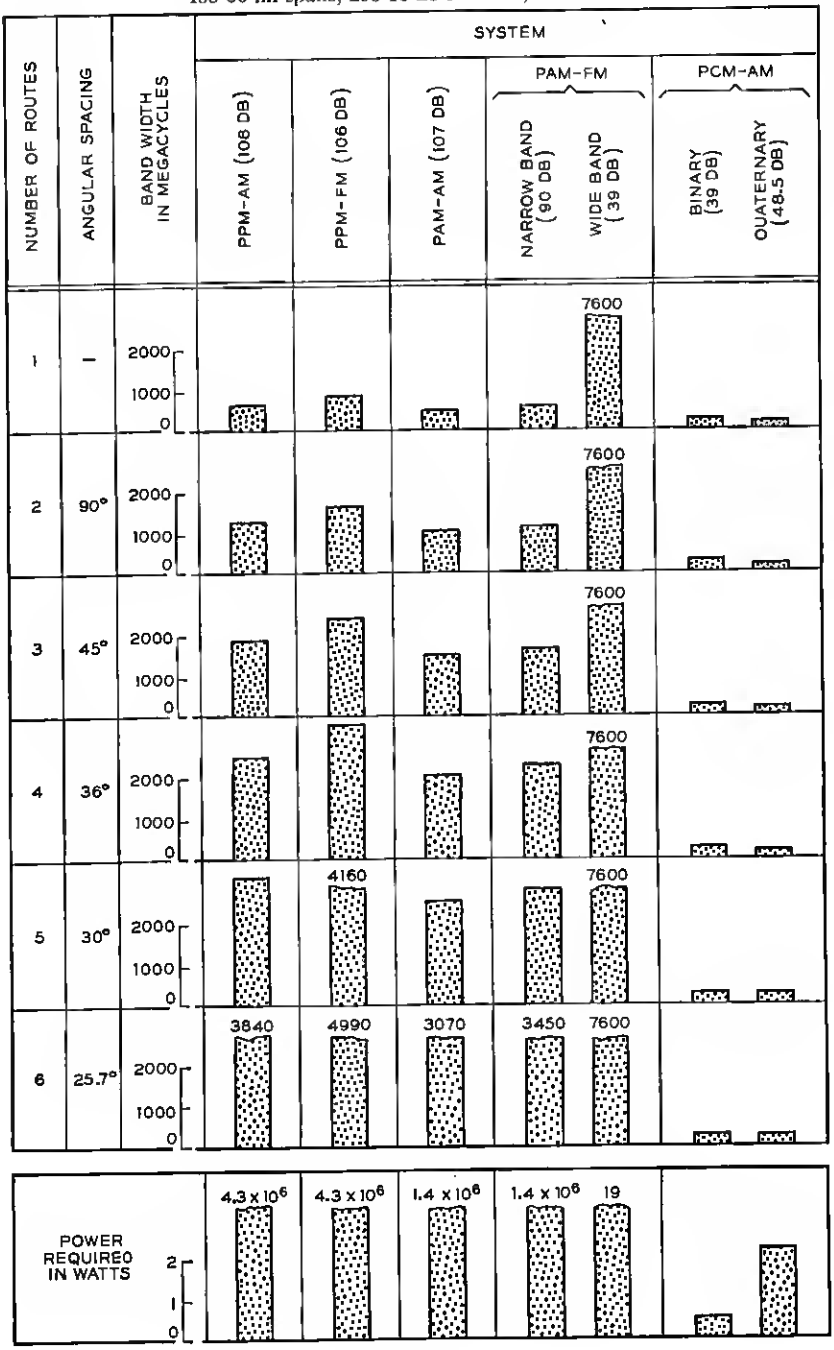


TABLE VII—*Concluded*

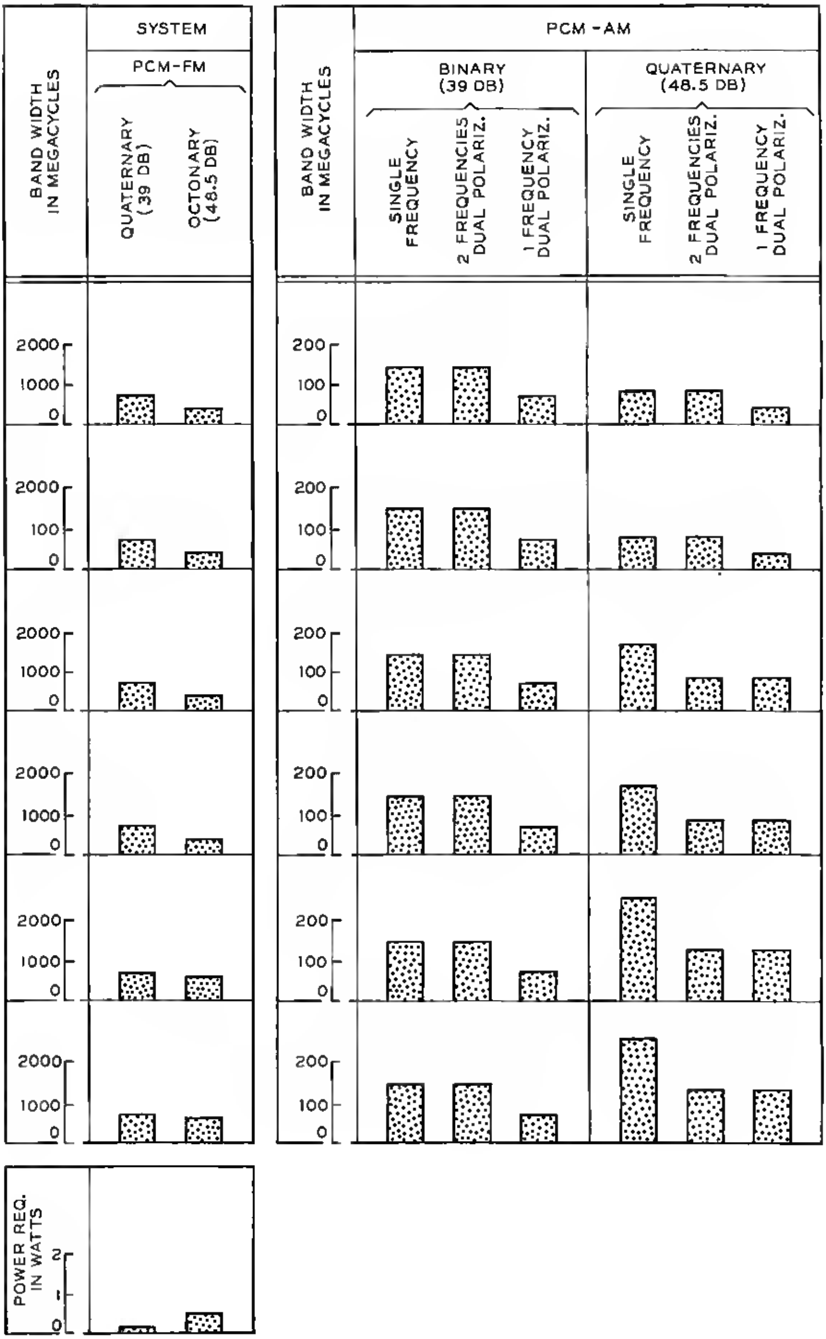
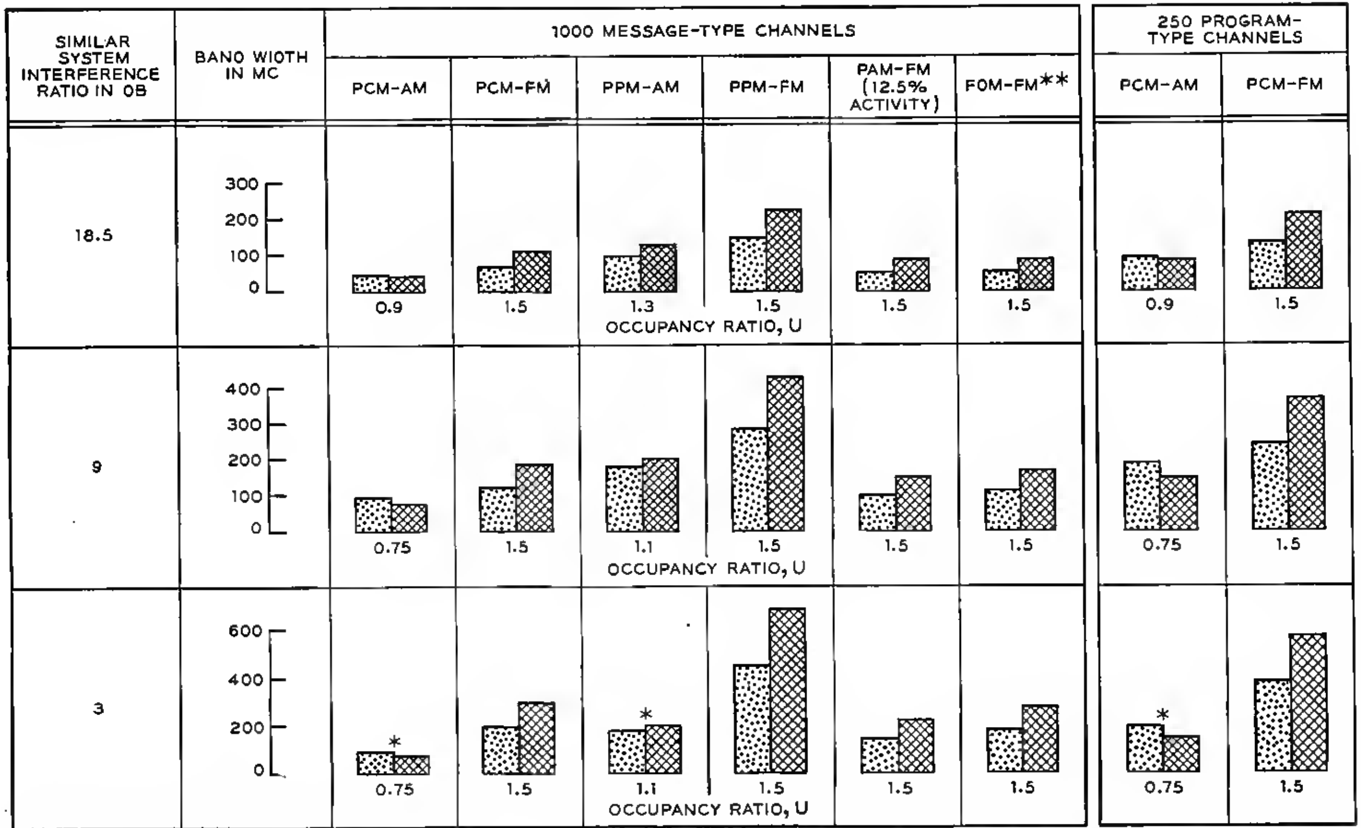


TABLE VIII

COMPARISONS OF BAND WIDTH AND FREQUENCY OCCUPANCY FOR SYSTEMS OF EQUAL RUGGEDNESS

Dotted bars show bandwidth; crosshatched bars show relative frequency occupancy.



*SUPPRESSED CARRIER, HOMODYNE DETECTION

**INTERFERENCE CALCULATED AS NOISE OF SAME POWER

These factors were multiplied by the product of bandwidth and number of frequencies to obtain the dotted bars in Table VI.

In regard to the "program grade" of circuit we must be more liberal in our allowance for guard space. Our estimates for the band spacing factor are:

System of Table VII	Factor U
PPM-AM	4
PPM-FM	4
PAM-AM	4
PAM-FM (narrow band)	4
PAM-FM (wide band)	3
PCM-AM (quaternary)	0.9
PCM-AM (binary)	0.75
PCM-FM (octonary)	1.5
PCM-FM (quaternary)	1.5

These factors were used to compute the dotted bars in Table VII.

If transmission on two polarizations can be accomplished with mutual cross-fire suppressed to a sufficient degree, half of the channels could be transmitted by each polarization, on the same frequency, thus halving the frequency occupancy. A probably unattainable cross-fire ratio seems necessary to meet the requirements in the non-regenerative systems, if we remember that the interference produced by cross-fire accumulates from span to span; but a suppression likely to be attainable, of the order of 15–20 db, makes this frequency saving feasible in the rugged systems such as binary PCM-AM or PCM-FM. The tables show entries for binary and quaternary PCM-AM, assuming dual polarization transmission.

If antennas could be improved to insure nearby discrimination ratios adequate to allow use of the same frequency in and out and west and east, the single-route occupancy would be halved again; with such a one-frequency repeater plan the occupancy in a congested area is not, however, always halved. Whenever the frequency requirements, as determined by the terminal antenna directivity, result in two or more frequencies, A, B, C . . . etc., there is no saving accruing from a one-frequency repeater plan, because two-frequency routes can be accommodated with no additional frequencies by suitably switching frequencies. It is only in the case of a system so rugged that the terminal antenna directivity permits a single frequency, A, to be used that the occupancy is reduced and it is then halved. With PCM of low base this is a possibility and the tables include entries for this case.

As to achieving antenna characteristics suitable for one-frequency operation, it may be noted that reflection from a heavy rainfall in front of the

antennas limits the attainable side-to-side ratio.[32] Reflection from aircraft may also impose a practical limitation. Spacing the antennas laterally (on two towers) would achieve freedom from these limitations. Another way of coping with the antenna discrimination obstacle is to use short spans in congested areas. This reduces the discrimination requirements particularly because fading is reduced by shortening the spans.

Conclusions as to Radio

Of the systems included in Table VI we find that, for six routes, binary PCM-AM, even without the potential frequency economy of dual polarization and/or single-frequency repeaters, has come close to being the most economical of frequency space; quaternary PCM-AM shows a slight advantage (which would be lost if the route spacing were less than fifteen degrees). Even without dual polarization or single-frequency repeaters, the binary PCM-AM occupancy is less, for more tban 3 routes, than the *occupancy* required by FDM whose *band width* is 4 kc per channel. Tbere is here an excellent illustration of the possibility of a net saving in frequency space through the use of tough wide-band systems.

The power requirements also favor the low-base PCM systems. It should be noted, in particular, that the linearity requirements with FDM demand that the tabulated power of 80 watts be a very light load on the repeaters.

Inspection of Table VII brings out the effectiveness of the coding principle if very high-grade channels are required. Only with PCM (of low base, as shown) are the occupancy and power requirements both within the practical realm. The non-PCM methods that achieve small occupancy, comparable with that of low-base PCM, all require colossal amounts of power. When the power requirement is reduced and the ruggedness increased by use of band width, the occupancy becomes, in turn, colossal. This is illustrated by the two entries for PAM-FM.

As route congestion increases without limit, any type of system that permits exchange between bandwidth and ruggedness will always achieve the minimum occupancy when bandwidth has been used to secure the degree of ruggedness that avoids multiplying the frequency assignments. Our studies have shown tbat, with the assumptions made, this result is valid for channels of message grade when the congestion has reached a degree that is by no means fantastic. We have accordingly prepared Table VIII in which the dotted bars show the bandwidths (taken from Fig. 9–19) of the various systems when their interference tolerances are alike and have values of 18.5, 9, and 3 db.[33] While these systems, having the same tolerance, all

[32] Measurements made at the BTL radio laboratory at Holmdel, N. J. indicate that this limit to side-to-side ratio is of the order of 85 to 90 db.

[33] The AM pulse systems are here assumed to achieve the 6 db increase in tolerance by suppressing the carrier.

fare alike in respect to frequency requirements imposed by antenna directivity, the bandwidth figures do not adequately reflect the merits of the systems. This is because the band-spacing factors are different and, in addition, only the regenerative systems can be expected to achieve the halving of occupancy accruing from dual polarization and from one-frequency routes. The crosshatched bars of Table VIII include the effect of multiplying by the estimated band-spacing factors shown beneath the bars. These band spacing factors are in some cases smaller than those previously tabulated for the less rugged systems of Tables IV and V. Only the PCM methods are shown for the case of very high-grade channels, since the non-PCM methods are so strikingly less effective here.

These conclusions depend for validity on the assumptions made and particularly on those concerning antennas, route disposition and fading, and apply when the converging systems *are of the same kind*. In a real situation, departures from the assumed conditions could markedly affect the conclusions. For instance, the meritorious showing of PCM in respect to efficient utilization of frequency space in the face of route congestion depends heavily on the assumption that all routes in the occupied space employ PCM. Any routes employing a modulation method that is highly vulnerable to interference like some of the narrower bandwidth methods would have to employ higher power to operate in the face of interference from the PCM routes. This higher power, concentrated in a narrower band, could destroy the PCM routes. In some cases it would obviously be impossible to assign values of power which permit the two kinds of routes to share the same frequency band.

Our calculations should be taken to illustrate the factors involved and the philosophy by which such problems may be approached rather than to find an unequivocally best system.

V. More About The Non-Simultaneous Load Advantage

The transmission advantage enjoyed by multiplexing many single sideband telephone channels in frequency division, discussed in the introduction, stems from several factors:

1. During the busiest period, only a small percentage (of the order of 12 to 15%) of the channels are actually transmitting speech ("talk spurts") at one time, on the average.

2. There are only a few *loud* talkers; the remaining ones range downward to a volume 35 to 40 db lower.

3. In the addition of the sideband voltages representing the talkers actually producing talk spurts, only a fraction of the grand maximum occurs often enough to be significant.

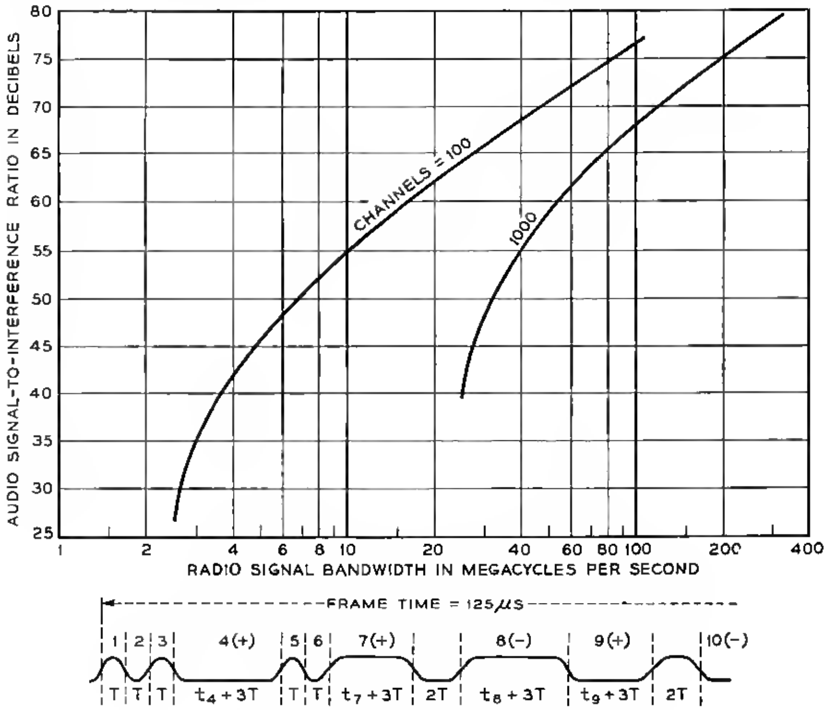
With frequency division all of these factors jointly contribute in a natural

and automatic manner to the low peak load ratings given in Table I. In time division, complicated instrumentation is needed to obtain such a low load rating (in time, now, not power capacity) and the saving is in bandwidth (time). Savings accruing from item (1) ahove are theoretically obtainable in all time-division systems (and, in fact, in nonmultiplexed multipair cable transmission systems) hy having automatic devices which skip tbe channels that are momentarily inactive and which advise the receiver of the skipping. It is possible also to henefit from items (2) and (3) ahove in systems which transmit a time interval to represent an amplitude. The amplitudes may be sent as absolute magnitudes together with a polarity indication. If this is done the channel time allotments actually required in a given multiplex frame appear piled up end to end, and many more channels can be handled than if provision were made for full amplitude on all. PPM is one such system, and PCM is another if the code symbols containing fewest digits are used to represent the smallest absolute magnitudes.

The use of instantaneous companding, which tends to make all talkers contribute equally to the system load, reduces the advantage represented by (2) above, but does not basically affect (1) which represents a substantial part of the total multiplex advantage.

It is illuminating to compute the performance of a pulse lengtb modulation system (PLM) employing the elastic time allotment and assuming that the load ratings of Table I apply. We imagine a system working on the principles illustrated in Fig. 23. There we assign a time $T(= 2/B)$ to each inactive channel. Active channels wbose ahsolute amplitudes are described by t, are assigned $t + 3T$ and those that are negative are preceded by a $2T$ pulse to designate that they are negative. If the interference is no greater tban marginal (9 db down) the receiver can distinguish between (a), the T intervals which count off the channels that are skipped and (b), the $2T$ polarity indications and (c), the $3T$ minimum signal intervals. The frame time of 125 microseconds must include the sum of these intervals plus Kt_0 wbere t_0 is the time shift for a full-load tone in a single channel which gives the required signal-to-noise ratio for the handwidth $B(= 2/T)$. The load rating factor is K, expressed as an amplitude ratio. The relations used to plot the two curves of Fig. 23 are sbown in tbe insert. Little or no instantaneous companding could be used to advantage so that a signal-to-interference ratio of 50 to 60 db would be required and for 1000 channels the handwidth would be between 30 and 50 mc, whicb is some two or three-fold less than in hinary PCM-AM, both systems being equally tolerant to a single source of CW interference. The elastic principle could presumahly be applied to PCM also to achieve a several-fold bandwidth reduction, but no experience has heen obtained with any of these elastic systems. While this paper has avoided for the most part questions of instru-

mentation, it should be pointed out that the elastic schemes tend to become complex apparatus-wise. If one chooses to discount this on the grounds



N = NUMBER OF CHANNELS

a = NUMBER OF ACTIVE CHANNELS

$$125 = N\left(1-\frac{a}{N}\right)T + N\frac{a}{N}\,3T + \frac{1}{2}N\frac{a}{N}\,2T + t_0K = \frac{2N}{B}\left(1+3\frac{a}{N}\right) + t_0K$$

$$t_0 = \frac{125-\frac{N}{B}\left(2+\frac{6a}{N}\right)}{K}$$

$$\text{SLICER ADVANTAGE} = 20\,\text{LOG}_{10}\,\frac{\pi}{2}\,\frac{t_0}{T} - 3\,\text{DB}$$

$$\frac{S}{I} = \text{RF PULSE TO INTERFERENCE RATIO} + 20\,\text{LOG}_{10}\,\frac{\pi}{2}\,\frac{t_0}{T} - 3$$

CALCULATED FOR: RF RATIO = 9 DB (MARGINAL)
N = 1000 K = 6.3 (16 DB)
N = 100 K = 2.83 (9 DB)
$\frac{a}{N} = \frac{1}{8}$

Fig. 23—Theoretical possibilities of exploiting non-simultaneous load advantage by an elastic PLM-AM system.

that future developments may resolve the complexity, there remains the objection that any system designed to take advantage of the multiplex load rating counts heavily on being used almost exclusively for conversational

speech under present operating procedures. The extensive use of telephone channels for nontelephone purposes is thus curtailed.

VI. Overload Distortion and Noise Threshold

In designing a microwave system for a large number of channels the power required to override noise may exceed the power capacity of available amplifiers. Also, the bandwidth may exceed the limit imposed by microwave transmission phenomena or circuit techniques. In either case, the remedy is to divide the channels into several groups of fewer channels and transmit the groups in adjacent narrower bands spaced by the proper factor U, and separable with filters for individual amplification, reshaping or regeneration. The power requirement falls off linearly with bandwidth. The filter problem for AM pulse transmission is considered in Section VII. The total frequency occupancy is no greater for this division since the same percentage "guard band" is involved if, in both cases, the neighboring, foreign signals are of the same kind as the wanted signals. In case the neighboring signals are of a different kind, the multiple band arrangement is in fact likely to represent a smaller occupancy because the occupancy is in general more sharply defined when made up of several narrower bands.

When considering a multiple group arrangement, it may be economical to provide for a substantial amount of common amplification prior to separation into the several bands which receive individual treatment. The nonlinearity of the common amplifier then sets a limit to the common amplification. Experiments bearing on this overload limit were made with the PCM equipment described by Meacham and Peterson.[12] Two- and eight-frequency groups were employed and the amplifier load was increased until the effects of distortion began to appear. The distortion was measured in terms of the maximum amount of CW interference which, when added to the amplifier output, resulted in no audible effect in the PCM channels. The right-hand part of Fig. 24 plots the results. For eight bands (six of which were not pulsed but were left on as unmodulated carriers) it is seen that the margin provided against CW interference begins to shrink rapidly when the single group load is 20 db below the output at which 1 db compression occurs. Tbe margin is completely used up (the channels begin to show noise) when the load is 13 db bigber. The left-hand part of Fig. 24 plots the manner in which the low level limitation (noise) was found to appear. Margin against CW interference shows a reduction for a pulse-to-noise ratio of 28 db and is completely used up at a ratio of 18 db.

The overload occurred in the 4000-mc power amplifier associated with the repeater, and the noise originated in the receiver.

In non-reshaping amplitude-modulated systems, the effect of compression

[12] Loc. cit.

occurring in the repeaters is cumulative. In microwave repeaters second-order distortion products fall outside the band and third-order distortion is likely to be predominant. We assume in what follows that the distortion arises solely from a cubic term. When the low-level gains of the repeaters are maintained equal to the preceding span losses, it can be shown that the single-frequency compression characteristic at the end of n spans is approxi-

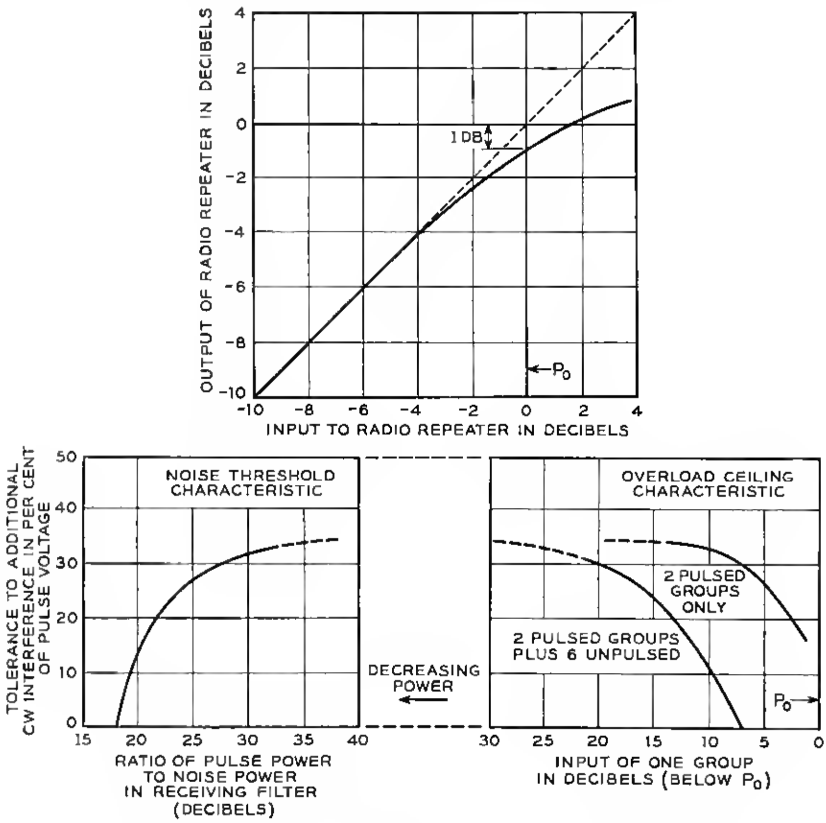


Fig. 24—Noise threshold and overload ceiling in frequency divided PCM groups.

mately the same as for one span but occurs at a power level 10 log n db lower. This approximation becomes more exact as the over-all distortion involved becomes less (as by lower input power). Fig. 25 shows a third order compression curve for one span and the resulting curves, obtained graphically, for 2, 4 and 10 spans. Examination of these curves shows that the curves are substantially the same as for the single span but displaced, 3, 6 and 10 db respectively. This is illustrated by line A, which intersects all of the curves at the same compression value (1.7 db). The points

marking the intersections are seen to be displaced from the intersection with the curve for one span by approximately 3, 6 and 10 db. If the phase of the repeaters is as linear as it must be in pulse systems, this single frequency characteristic can be applied for the entire signal band as if it resulted from a single source of third-order distortion. The effect of this distortion is

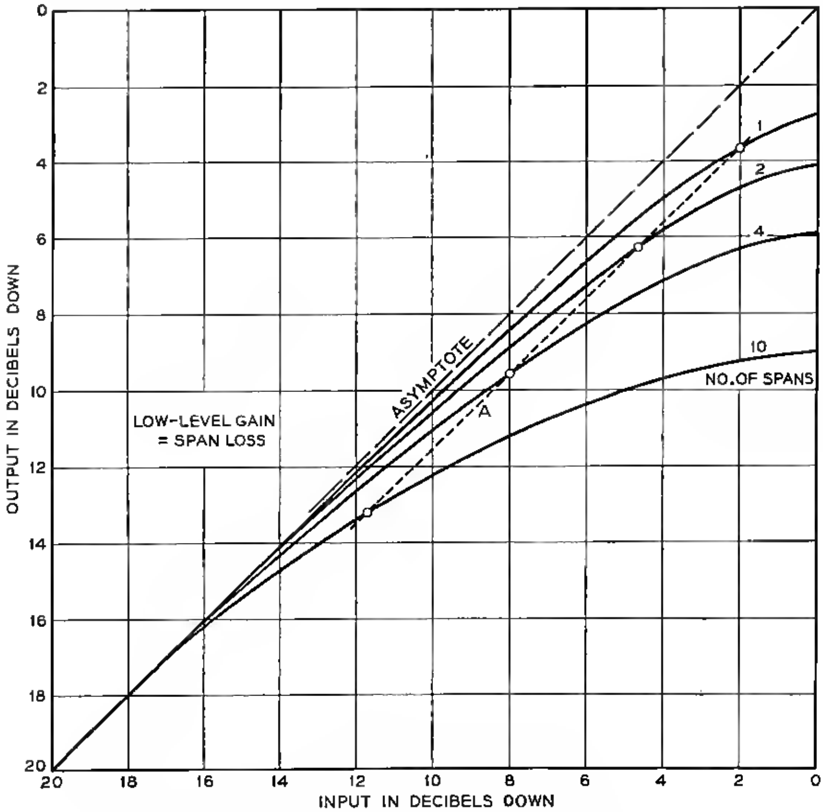


Fig. 25—Overload characteristics of multi-repeater systems.

serious in multiband PCM repeaters, as illustrated by the measurements on Fig. 24, but is, generally speaking, less important in single band pulse repeaters. For instance[34] PPM-AM pulses and binary PCM-AM pulses might operate on the flat part of curve 10 (Fig. 25). With PCM of higher

[34] On the grounds that pulse slicers themselves include the compression function to a high degree, one might not see the harm of compression in repeaters. If all of the compression occurred *after* the noise had been acquired there would be no fundamental compression penalty in slicing pulse systems. The penalty comes about because as the pulses progress from span to span they shrink and become more vulnerable to noise.

base as well as with PAM, the repeater loading would have to be sharply reduced, however.

More power on all spans could be obtained by making the repeater gain greater than the span loss. This very quickly defeats its purpose, however, because the excess low-level gain raises low-level noise between pulses to a high level status as it progresses from span to span.

Reshaping of AM pulses (and of course regeneration in PCM-AM) at all repeaters avoids the cumulative effect of compression by permitting the repeater gain to be greater than the span loss by the amount of compression on one span.

When a signal is transmitted by FM, the phase curve of the transmission circuit plays a role somewhat analogous to the amplitude characteristic of an AM system. The correspondence is not complete, however, for we find that modulation products arising from even-order phase distortion as well as from odd fall in tbe signal band even though the FM band is located in a frequency range very high compared with the baseband width. For amplitude modulated signals in the baseband, we can replace the FM pbase distortion effects by an equivalent non-linear baseband amplifier characteristic which has the same shape with respect to zero voltage input as the phase characteristic has with respect to the midband frequency of the FM range. If the distortion is small, the square and cube law approximations obtained by expanding the phase-shift function about the mid frequency may be applied as in conventional multichannel cross-modulation theory.[35] We shall not here attempt to discuss the accumulation of phase distortion in a multi-repeater FM system.

VII. Pulses, Spectra, and Filters

In this section, we will consider: (1) pulse shapes in relation to the particular pulse modulation method employing them, (2) the shaping filters by which they may be obtained and (3) the transmitting and receiving filters employed in systems comprising a multiplicity of adjacent frequency bands each carrying pulse signals.

Column A of Fig. 26 shows various pulse shapes which can be approximated (with the exception of shapes 8 and 9) by fairly simple circuits, both in the baseband and radio spectrum. Pulse 1 is an "unshaped" rectangular pulse. A good approximation to it can be obtained in wide-band circuits accommodating the extensive spectrum it possesses, i.e., in circuits having rise and decay times short compared with the duration T_0. Such a pulse when transmitted through Gaussian filters of the various widths shown in

[35] W. R. Bennett, "Cross-Modulation in Multichannel Amplifiers" *Bell System Technical Journal*, Vol. 19, pp. 587–610, October, 1940.

column C emerges with smooth transitions as shown in 2, 3 and 4. These pulses rise and fall in a nearly sinusoidal manner. The width between half-amplitude points is T_0. Shortening the rectangular pulse ("curbing") and narrowing the shaping filter can be made to result in pulses 5 and 6 which have the same width between, say, 3% points (at t_1) as pulse 4.

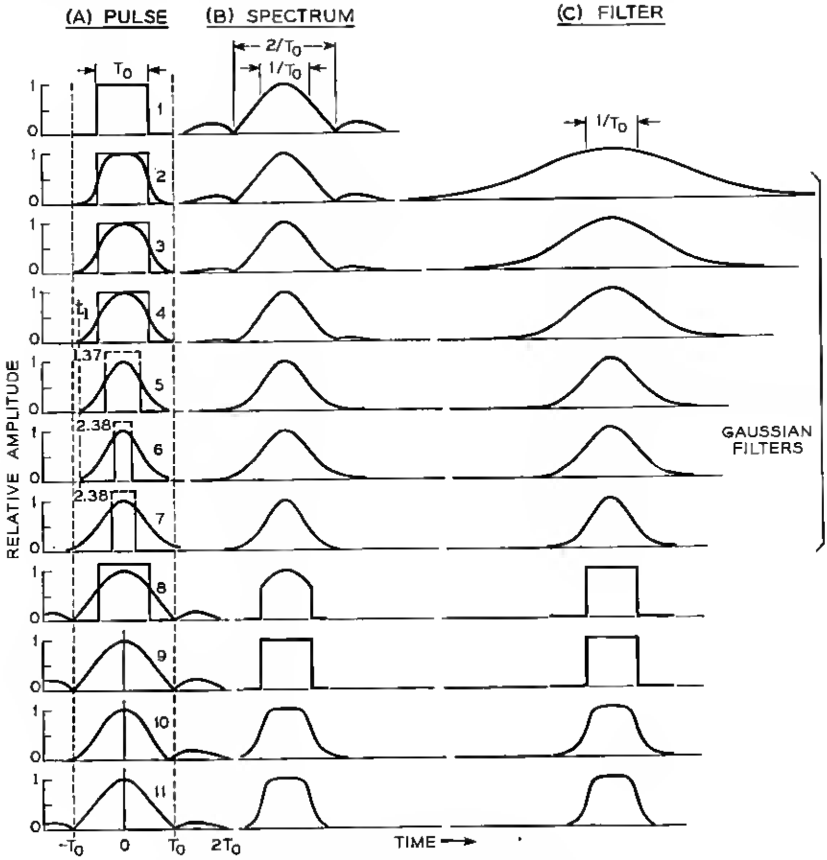


Fig. 26—Typical pulses and their spectra.

Pulses 5 and 6 are then shorter than 4 between half-amplitude points. If the half-amplitude width is made the same as in pulse 4 the width between lower amplitude points is greater than in pulse 4. This is illustrated in pulse 7.

Gaussian filters[36] as defined here are naturally linear phase networks

[36] Gaussian filters are networks whose transfer admittance follows the error law as a function of frequency. A decibel plot of a bandpass Gaussian filter is accordingly a parabola in shape.

and we have assumed linear phase in computing pulses 2 to 7. A good approximation to the Gaussian filter can be obtained both as to phase and amplitude with a number of tuned circuits in tandem, coupled through buffers. A fair approximation can also be obtained by combining a 3- or 4-section maximally flat filter[37] with a tuned circuit through a buffer.

Rectangular or near-rectangular shaping filters produce pulses with overshoot as shown by pulses 8 to 11. The filter corresponding to pulses 8 and 9 is assumed to have rectangular shape and linear phase. Filters of this sort have no simple approximation in practice and are included for comparison with filters 10 and 11 which are made up of simple maximally flat networks. In pulses 9, 10 and 11 the "unshaped pulse" is assumed to he very narrow and of amplitude sufficient to yield pulses of the heights shown.

Let us now regard these pulses as received pulses and compare them in respect to shape for use in various kinds of pulse systems.

PPM. In PPM the pulses may occupy any time position in the assigned interval and so the tails of pulses 8 to 11 may "crosstalk" into time assigned to an adjacent channel. To allow guard time for the train of tails or to design for satisfactory operation in the presence of the tails is uneconomical of frequency space. It follows that pulses which are more definitely bounded in time such as those obtained with Gaussian filters are more desirahle and likely to be more economical of frequency space in general despite their wider spectrum.

In PPM where the trailing (or leading) edge of a pulse is used to convey the information a flat top pulse such as pulse 2 is no better than one in which the flat portion is ahsent and the two transitions brought together.[38] The latter pulse would, in fact, be superior since more time would then be available for additional channels or for greater swing.

We are thus led to conclude that one of the pulses in the 4 to 6 group is the preferred shape for PPM. We chose pulse 4 in our illustrative calculations and defined bandwidth as $2/T_0$, but pulses 5 or 6 would have given substantially the same results.

PAM. In PAM the pulses occur at standardized, regular times so that if pulse 9 were used the accompanying tails, which disappear completely at instants T_0, $2T_0$, etc., from the pulse peak, need not theoretically produce crosstalk between channels if the channels are spaced T_0 and the pulse amplitudes are measured *instantaneously* at the time the nulls occur. As a practical matter both the precise pulse shape and the instantaneous measure-

[37] W. W. Mumford, "Maximally-Flat Filters in Wave Guide," *Bell Sys. Tech. Jl.*, Vol. 27, October, 1948, pp. 684–713.

[38] Such a pulse would look like pulse 4 if the latter were shrunk to occupy 0.6 of the time shown in the plot. The spectrum would accordingly be that of pulse 4 expanded by the factor 1.7 but would not include more significant band width than is necessary to form pulse 2 as shown. This deduction follows from the fact that the rise time of pulse 2 is the same with or without the flat top.

ments at precise instants are probably not realizable to a degree which would keep the crosstalk within tolerable limits, so that one of the smooth pulse shapes is preferred. Pulse 4 with a spacing of T_0 is feasible from the sampling precision point of view but a spacing of $2T_0$ provides margin against crosstalk arising from small imperfections in any realizable approximation to the theoretical pulse.

It is to he noted that, if an instantaneous sample is taken of a PAM pulse, the measured magnitude is affected directly hy the instantaneous value of noise present in the entire hand occupied by the pulse. No frequency selectivity can be applied afterward to remove the influence of any part of the noise band hecause the error, even though caused hy wide-band components, is exactly the same as could have heen produced by a uniquely determined wave wholly confined to the signal hand itself. The hest signal-to-noise ratio obtainable with instantaneous sampling is that associated with minimum bandwidth for the pulse (i.e., pulse 9) and the corresponding maximum stringency of synchronization requirements on the sampling and pulse distortion. The same signal-to-noise ratio can, however, be approached with a wider band provided that we allow a finite segment of the received pulse to enter the channel filter. An averaging out of higher-frequency disturhances produced by wide-hand noise is thus attained.

PCM. In PCM a short sample taken near the center of a pulse serves to determine correctly the presence or absence of a pulse even in the presence of interference at or near the hreaking point of the slicer. Thus, pulse 4 may be used with a spacing of T_0, and if a gate pulse 25% of T_0 is used, it need not be aligned with an inordinate precision to obtain good operation.[39] Greater tolerance in the matter of sampling would be obtained with pulse 2 but the frequency extravagance could scarcely be countenanced. As stated we assume pulse 4 in our PCM bandwidth curves but employ pulse 11 in Tables VI–VIII. Use of pulse 11 is a frequency conservation measure that seems feasihle only with PCM and is attractive only with binary PCM.

Optimum Distribution of Selectivity Between Transmitting and Receiving Filters

In a regenerative repeater system both the receiving and transmitting filters may be Gaussian without suffering cumulative narrowing of the system bandwidth since each span commences with a freshly shaped pulse. In this case, the transmitting filter of one repeater and the receiving filter of the succeeding repeater comhine, as Gaussian filters do, to make another Gaussian filter. The resulting pulse may he one of the series 2 to 6 of Fig. 26. On the assumption that one of these shapes is desired and that the trans-

[39] This is the pulse shape approximated in the experimental system described by Meacham and Peterson (loc. cit.).

mitting and receiving filters are to be Gaussian, a problem arises as to how to divide the total selectivity (in column C) between them. If most of the pulse shaping is done at the transmitter, the Gaussian receiving filter must be extremely broad, with the result that discrimination between pulses in adjacent bands is poor and the bands must be spaced widely in order to keep cross-fire down. If, on the other hand, all of the shaping is done at the receiver the wide spectrum of the unshaped transmitted pulse spills over into neighboring bands unless the bands are widely spaced. Clearly, an optimum proportioning of selectivity exists and it is interesting and enlightening to analyze this problem. Such an analysis was made for pulses 4, 5 and 6. This analysis pertains only to crossfire and not to signal-to-noise ratio as influenced by curbing (shortening of the rectangular pulses) and by the division of selectivity between transmitting and receiving filters. Wide receiving filters accept more noise and narrow ones may prevent the transmitted pulse from attaining full height in the receiving filter output if curbing is used. If the curbing is pronounced, as in pulses 6 to 11, amplification may have to follow the transmitting filter to establish the desired transmitted power level. For divisions of selectivity close to the optimum, the receiving filter selectivity appreciably reduces the transmitted pulse height in the case of pulse 5 and seriously reduces it in the case of pulse 6.

Crossfire from a pulse in an unwanted band appears as a transient in the wanted band. In some circumstances, this transient has peaks which occur while the crossfiring pulse is rising and falling and has a minimum between which sometimes dips below the level fixed with the steady-state discrimination to the crossfiring carrier. If the pulses in the crossfiring band are synchronized with those in the wanted band as they might be in PAM and PCM only the minimum, central, crossfire might be significant. If, as in PPM, the pulses cannot be synchronized, the peak crossfire is significant. Curves for two values of band separation are shown in Fig. 27, one appropriate to yield minimum crossfire in the 25 to 35 db range and the other to yield peak crossfire in that range. This is the range that is sufficient for binary PCM. The steady state discrimination is also shown. We conclude from this study that pulses 4 or 5 are about equally good in respect to minimum central crossfire and that pulse 4 is slightly preferable in that the trough and the crest are more symmetrical. For PCM in which the pulse spacing is made equal to T_0 this symmetry means that there is the same margin for misalignment of the gating pulse, as regards correctly interpreting a space or a mark. Pulses 5 and 6 appear to be about equally good in respect to peak crossfire but both (and particularly pulse 6) incur a signal-to-noise penalty because the receiving filter does not permit the transmitted pulse to attain full height.

In practice, the approximations to Gaussian filters have shown worse

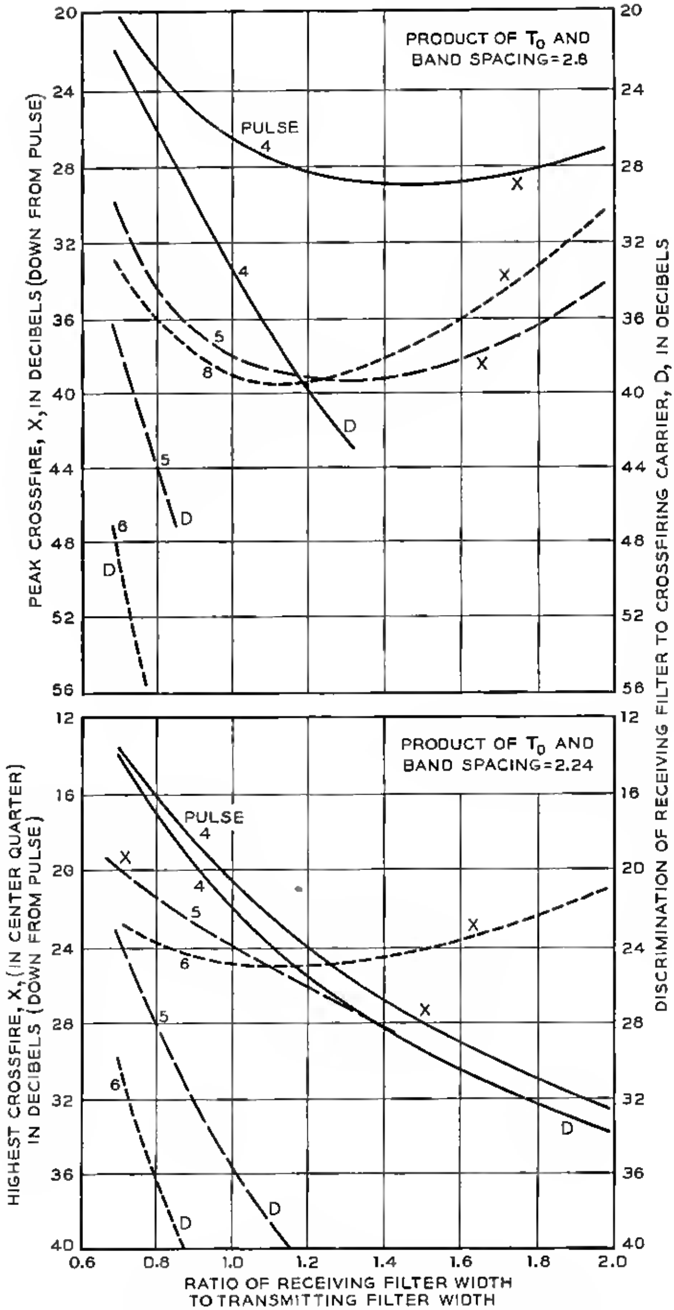


Fig. 27—Crossfire between frequency divided pulse groups.

pulse crossfire than the curves predict. This is particularly so for cases in which the curves show crossfire 30 or more decibels down. Approximations usually possess less rapidly falling attenuation skirts and possess phase distortion, both of which prevent realization of the calculated crossfire values.

Because regeneration (or reshaping) *permits* the use of Gaussian receiving filters, it does not follow that *flat-topped* filters are inferior as receiving filters. Calculations were made for maximally flat receiving filters of about the same overall complexity as was involved in the Gaussian approximations. They showed that when the transmitted pulse has the shape 4 and the flat filter is scaled to transmit such a pulse without much distortion, values of peak crossfire of the order of 30 db can be obtained when the product of band spacing and T_0 is 2.8. It was also found that the crossfire in that case consists of a single peak (not unlike the main pulse) nearly coincident in time with the crossfiring pulse. Our Gaussian approximations gave peak crossfire of this same order, for band spacing times $T_0 = 2.8$. The maximally flat receiving filter accepts roughly twice the noise power accepted by the optimum Gaussian filter, so the favor remains with the Gaussian filter and pulses 4 or 5.

The main conclusion from all of this is that, if smooth pulses, like numbers 4 or 5, are employed, band spacings of the order of $2.8/T_0$ (perhaps $2.5/T_0$) can be used with crossfire entirely suitable for binary PCM, as well as for PPM systems with sufficient swing ratio. Larger spacings would be required for PCM using multi-valued digits, and for PAM.

As mentioned earlier, the use of pulses 10 or 11 spaced by T_0 is possible in binary PCM, with small penalty, if very short accurately aligned gate pulses are used. The spectrum of these pulses is more sharply defined and includes a band only slightly wider than $1/T_0$. Rectangular receiving filters of that width could be used side-by-side so that the band spacing would be only slightly greater than $1/T_0$. This is the "theoretical minimum" and in telegraph parlance would be specified as a band spacing of twice the dot frequency.

Pulse 10 results from transmitting a very short pulse through a 4-section maximally flat filter whose response is shown in Column C. The phase distortion characteristics of such a filter produces asymmetry in the pulse. Pulse 11 is produced by the filter shown, assuming that the distortion is corrected. Most of the pulse shaping is assumed to reside in the transmitting filter. The assumed receiving filter is a 4-section maximally flat filter, and therefore has the shape of the filter shown for pulse 10, but is about 30% wider than shown. When two such bands are spaced $1.5/T_0$ the maximum crossfire is about 26 db down.

With shaping and receiving filters of reasonable complexity a band spacing

of $1.5/T_0$ to $1.7/T_0$ can be expected to have satisfactorily small crossfire for binary PCM. Pulse 11 and a spacing of $1.5/T_0$ were assumed for binary PCM-AM in Tables VI to VIII.

Figure 26 shows the envelopes of r.f. pulses produced by passing flat-topped r.f. pulses or r.f. *spikes* through r.f. filters. These envelopes are the baseband shapes produced by wide-band envelope detectors. If baseband pulses are shaped by baseband filters the resulting pulses are the same as shown for pulses 1 to 7, but for pulses 8 to 11 the tails turn out to be overshoots passing smoothly through zero instead of reaching zero cusp-wise. If these pulses are used to modulate the amplitude of a carrier in a product modulator, the cusps in the envelope are produced as shown, but if they are used to modulate the frequency of a carrier the baseband pulses produced by frequency detection retain their smooth transition through zero. In PAM-FM relatively wide gate pulses could be centered at time zero, T_0, $2T_0$, etc., and the inter-pulse crosstalk would be partially balanced out by partial cancellation of positive and negative contributions. By the use of biases in the AM case a similar result could be obtained. Our tables, assuming pulse 4 spaced $2T_0$, do not reflect this possibility of operation.

Delay Line Balancing

Techniques have been developed[4, 40] in which the received pulse train is split into two or more branches, after detection to the baseband, and recombined with suitable delays, attenuations and polarity reversals. Such a procedure is effective in reducing the pulse tails or hangover and its use has been especially valuable in experimental PAM and PAM-FM systems. While this device may be regarded as a kind of phase and amplitude equalizer (comprising as it does only linear, passive elements) the result may be a pulse shape slightly more desirable than those obtained from simple but "ideal" networks, shown in Fig. 26. Our judgment that pulses of shape 4, spaced $2T_0$, should be used in PAM-FM may be slightly pessimistic if this kind of balancing is used.

More significant reductions of inter-pulse interference may be sought by the method suggested by MacColl[41] (which is more than "equalizing") but this method, like the PCM method of Appendix III soon makes preposterous demands on the transmission medium and upon the transmitted power.

VIII. Transmission Over Metallic Circuits

In radio relay transmission we have assumed a span length of 30 miles and have assumed span losses in keeping with the microwave antenna art

[4] V. D. Landon, loc. cit.

[40] W. D. Boothroyd and E. M. Creamer, Jr., "A Time Division Multiplexing System," Paper presented at winter general meeting, *A.I.E.E.*, New York, Jan. 31, 1949.

[41] *U. S. Patent No.* 2,056, 284 Oct. 6, 1936 issued to L. A. MacColl.

and the relatively meager propagation experience now available. The high cost of the towers, power facilities and access roads involved in repeaters as we know them points to the desirability of few repeaters and long spans. Topography usually permits spans of a few tens of miles without requiring towers of excessive height. Very much longer spans can rarely be had without tremendous towers and are questionable because of the increase of fading depth with distance.

In wave guide (or other metallic conductor) transmission entirely different considerations apply and we will discuss some electrical relationships which seem significant in this case.

Let us consider a microwave repeater having a noise figure NF and a power capacity PC. The overload characteristic, together with the amount of nonlinear distortion that the signal can stand, determines the maximum output power. This maximum power is the power capacity. These two characteristics, PC and NF, thus determine the amount of attenuation that may be introduced between the transmitting half of a repeater regarded now as a transmitting terminal and the receiving half regarded as a receiving terminal. This amount of attenuation expressed in decibels, which we will designate as M, is available to be used up by the loss of one span plus accumulation of noise from n repeaters and may be regarded as a figure of merit of the repeater.

Five different relationships apply as follows:

$$\text{AM Systems: } M = \text{span loss}_{db} + 20 \log n \tag{1}$$

$$\text{FM Systems: } M = \text{span loss}_{db} + 10 \log n \tag{2}$$

PPM-AM Systems with

$$\text{reshaping repeaters: } M = \text{span loss}_{db} + 5 \log n \tag{3}$$

$$\text{Band increased } \sqrt[2]{n} \text{ referred to (1)}$$

FM Systems with

$$\text{limiting repeaters: } M = \text{span loss}_{db} + 3.33 \log n \tag{4}$$

$$\text{Band increased } \sqrt[3]{n} \text{ referred to (2)}$$

PCM Transmission with

$$\text{regenerative repeaters: } M = \text{span loss}_{db} + \text{zero} \tag{5}$$

In (1) the 20 log n term includes 10 log n for noise accumulation plus 10 log n for cumulative compression. In microwave amplifiers only odd-order terms contribute to the distortion and the third order term predominates for moderate degrees of overload. This results in the well-known compression characteristic such as appears in Figs. 24 and 25 previously discussed. A significant approximation for the over-all compression when

n such amplifiers are connected in tandem is that the compression characteristic is the same as with one amplifier but occurs at outputs 10 log n db lower. Thus, the power level must be reduced 10 log n db and this penalty accrues over and above the noise accumulation penalty.

In (2) only noise accumulation occurs.

In (3) and (4) it is assumed that minimum power conditions are attained and the operation has reached the straight part of the minimum (marginal) power curves. Without reshaping, the system must be powered so that, at the final repeater, the accumulation of noise does not exceed the marginal value. With reshaping at each of the n repeaters each span may be marginal. Making each span marginal with the same bandwidth would be accomplished with 10 log n db less power and would make the signal-to-noise ratio 10 log n db lower. This can be made up by using more bandwidth. In marginal PPM-AM the signal-to-noise ratio improvement occurs at the rate of 20 db per decade of bandwidth and thus the bandwidth must be increased by $\sqrt[2]{n}$. This requires, to keep the operation marginal, an increase in power of 10 log $n^{1/2}$ = 5 log n db. In the case of marginal FM, signal-to-noise ratio is improved at the rate of 30 db per decade, and the bandwidth must accordingly be increased by $\sqrt[3]{n}$. To keep the operation marginal, the power must be increased 10 log $n^{1/3}$ = 3.33 log n db. The entries in Tables II and III invoke these relationships. There, n may be thought of as having values 1, 5 or 133.

Equation (5) reflects the fact that where PCM regenerative repeaters are employed no accumulation of noise occurs with number of spans.

With metallic conductors, the span loss in decibels is proportional to the length of span. If A denotes the span loss in decibels per mile and S denotes span length in miles, the circuit length $L = nS$ and

$$\mathrm{M} = \frac{AL}{n} + x \log n \tag{6}$$

or,

$$\mathrm{L} = \frac{n}{A} M - \frac{nx}{A} \log n \tag{7}$$

where x is the appropriate coefficient, 20, 10. 5 or 3.33. In this expression there is an optimum value of n corresponding to a maximum value of circuit length L. Figure 28 is a plot of circuit length for $x = 20$ (Eq. 1), showing the maxima. Figures 29 to 32 show the optimum values of n and the resulting maximum circuit lengths for each of the relations expressed in equations (1), (2), (3), (4).

Considerations affecting transmission over metallic circuits are different from those affecting radio relay in at least the following four ways:

1. Interference from other routes substantially vanishes with coaxial and

wave-guide conductors. This diminishes the premium on ruggedness provided sufficient power is available so that ruggedness with respect to noise is not critical.

2. Since there is no fading and all spans can be of approximately equal length, all spans will possess the same loss, approximately. This situation

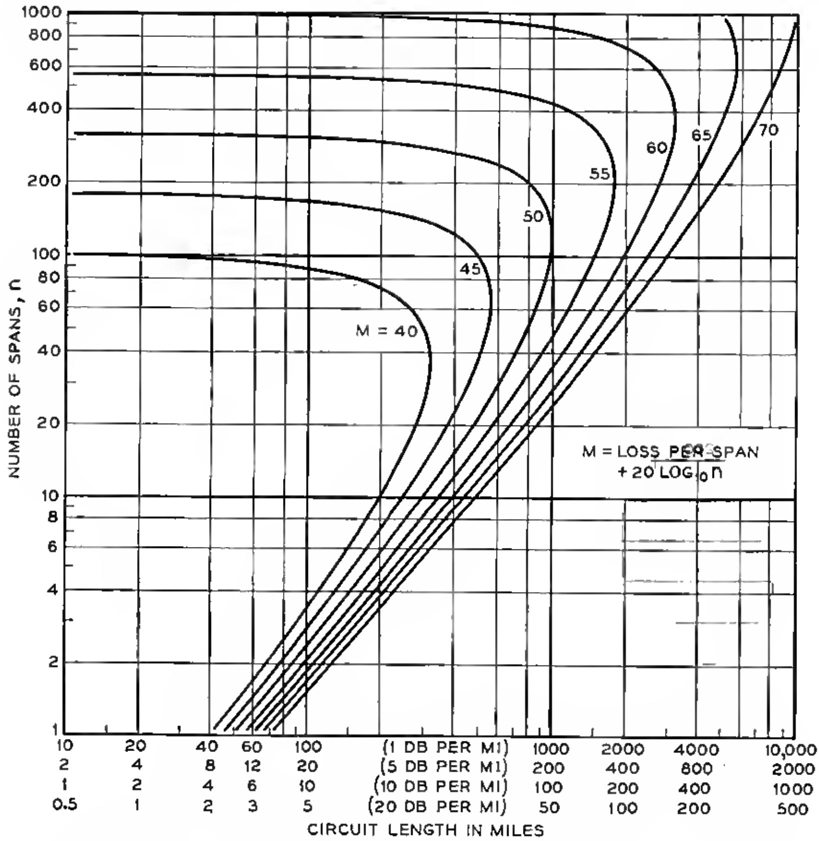


Fig. 28—Variation of circuit length with number of repeater sections in an AM system with fixed power capacity and noise figure.

is favorable to all systems but is most favorable to PCM, which gets no credit for low loss spans.

3. In the case of wave guide, frequency space may be much less precious than in coaxial or radio relay transmission.

4. There is a possibility that many small repeaters should replace the few higher powered repeaters used in radio relay.

These different considerations may lead to a different evaluation of the

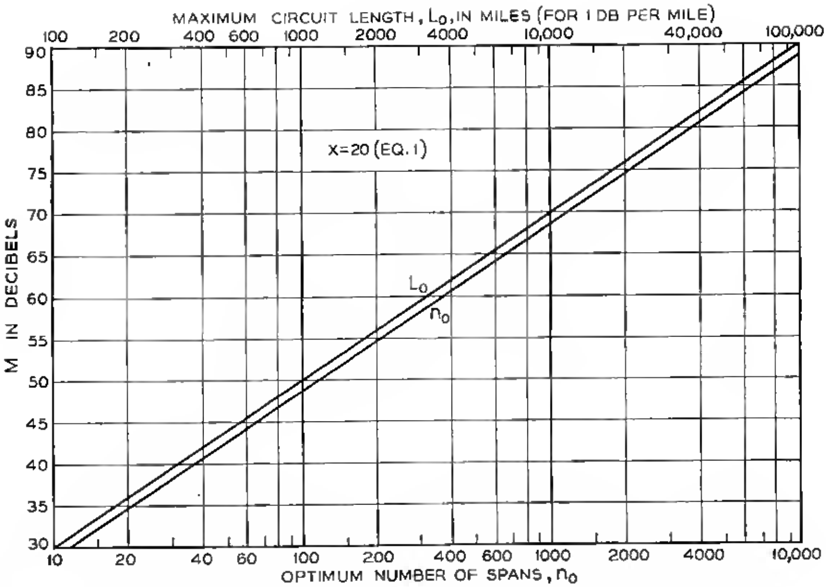


Fig. 29—Optimum number of repeater sections and maximum circuit length for metallic AM system with fixed power capacity and noise figure.

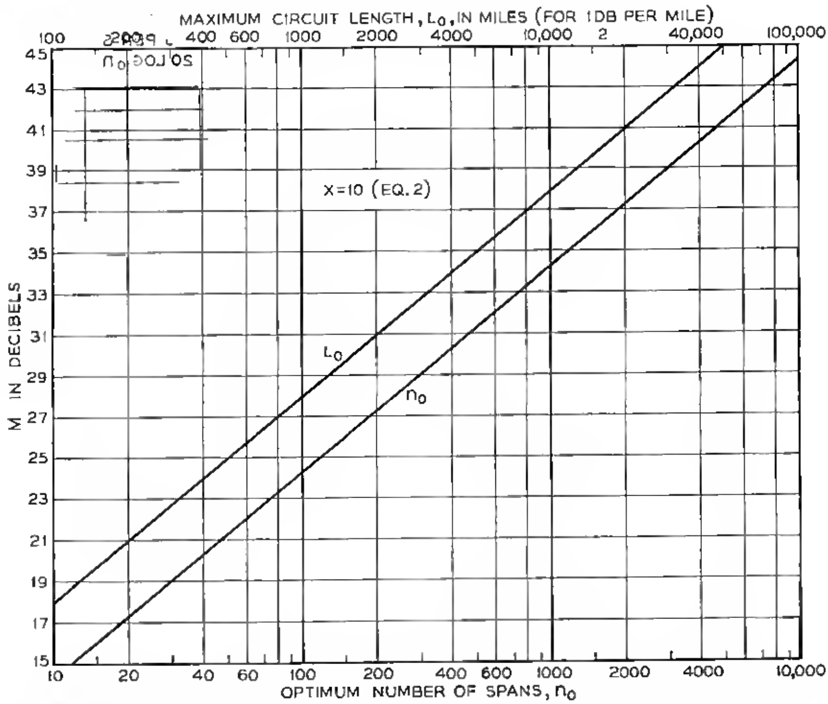


Fig. 30—Optimum number of repeater sections and maximum circuit length for metallic FM system with limiting only at end of system.

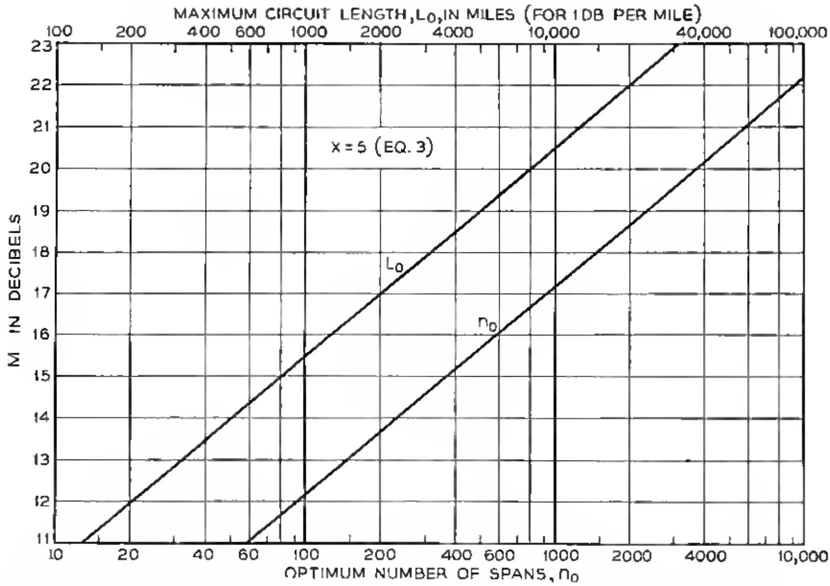


Fig. 31—Optimum number of repeater sections and maximum circuit length for metallic PPM-AM system with reshaping at every repeater.

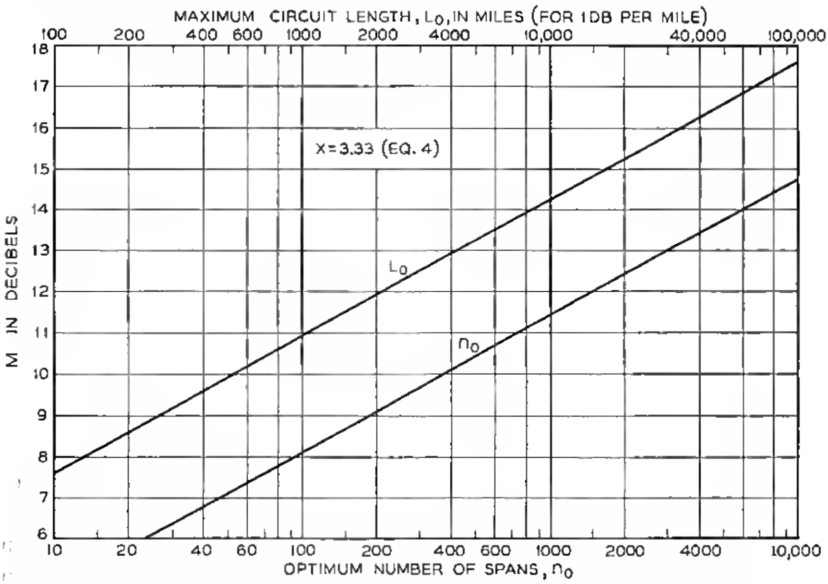


Fig. 32—Optimum number of repeater sections and maximum circuit length for metallic FM system with limiting at every repeater.

modulation methods discussed in this paper. We will not attempt to make such a re-evaluation here.

It is of interest now to return to radio relay transmission and examine the relations derived for metallic conductors, but now assuming that the span attenuation is that associated with an inverse k-power of distance law ($k = 2$ for free space attenuation). If we use the symbol E to denote the excess power capacity (in decibels) of the repeater over that required for a unit span of, say, one mile, we get the relation

$$10k \log L = E + (10k - x) \log n \tag{8}$$

where $x = 20, 10, 5, 3.33, 0$ for the cases described by equations (1), (2), (3), (4), (5) respectively. The equation shows no optimum number of spans

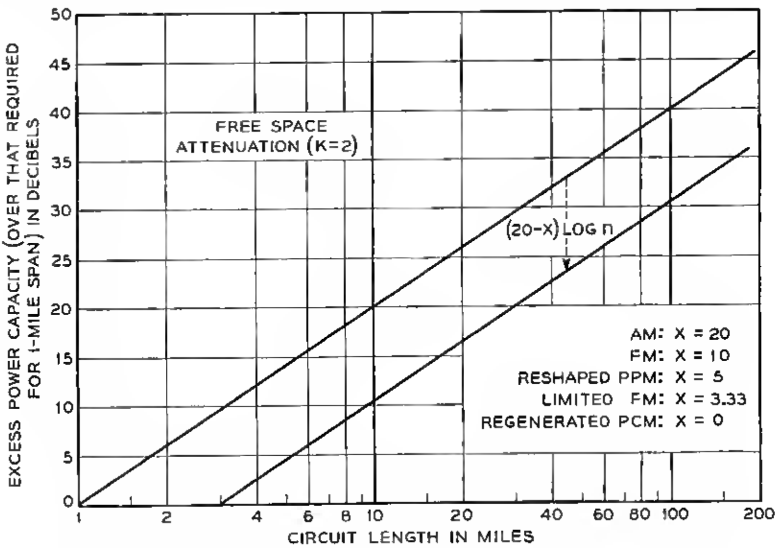


Fig. 33—Relation between circuit length, power, and number of repeaters in radio relay systems.

corresponding to a maximum circuit length. It also shows that when x is less than $10k$ the circuit length can be increased indefinitely by adding spans although the spans become shorter with increased circuit length. When $x = 10k$ the circuit length can not be increased beyond the maximum single span, i.e., it depends solely upon E and is not affected by the number of spans. If x is greater than $10k$ the circuit length again cannot be increased beyond the maximum single span and is reduced by employing more than one span. This last case does not occur for free space attenuation. In Fig. 33 is plotted the relationship between L, E and n for the free space attenuation law ($k = 2$). The curve passing through zero decibels excess power capacity at one mile circuit length applies to one span for any value

of x or to any number of spans when $x = 20$. In other words the maximum circuit length for $x = 20$ is the length of span corresponding to the excess power capacity as noted above for $x = 10k$. For all smaller values of x any circuit length can be achieved with any value of excess power capacity if a sufficient number of spans is employed. The number of spans required for a given circuit length is obtained by moving the curve downward until it intersects the desired length at the appropriate excess power ordinate, and equating $(20 - x) \log n$ to the downward shift in decibels.

Notwithstanding the present radio outlook in which large towers and antennas seem indicated, it is of interest to imagine small repeaters powered for a one-mile span, say. Using FM with limiting at every repeater, a 100-mile circuit could be obtained with 250 repeaters spaced 0.4 miles. This result comes from Fig. 33 with excess power = zero db and $x = 3.33$. A difficulty with such a case might be multiple paths produced by one repeater output overreaching into other spans.

The inverse k power attenuation does not accurately describe propagation over long spans; fading then occurs and is greater for long spans than short spans. This introduces a term in the span loss similar to that of the metallic conductor case in which the span loss is proportional to span length.

IX. Conclusions

We have, in this paper, examined some of the relations governing the exchange of bandwidth for advantages in transmission that grow out of the liberal use of bandwidth. While we have not dealt specifically with the instrumentation involved in the application of the various exchange methods, we have taken cognizance of certain basic obstacles in circuit design such as overload distortion, phase distortion and discrimination characteristics of selective networks and the limitations of microwave antennas. Not having, in most cases, a wealth of experience bearing on the manner in which these obstacles affect the transmission problem, we have been obliged to estimate their effect in many cases. Considerable unreliability in these estimates would not, however, much affect the broad purpose of the paper. The economic factor that is involved in achieving reliable operation of apparatus has been largely ignored, although methods that seem to lead to fantastic instrumentation have not been given much attention.

Ruggedness of the transmitted signal, which is obtained at the cost of increased *bandwidth* can, properly handled, be made to conserve *frequency occupancy* in two ways: (a) ruggedness reduces the required "guard space" between one band and neighboring bands carrying other signals; (b) ruggedness reduces the multiplication of frequency assignments necessary in congested radio route situations.

For wave guide systems, the inter-route interference problem arising from

route congestion disappears but ruggedness is still a valuable feature. As to PCM, regeneration is an outstanding asset applicable also in wave guide transmission. In the case of very high-grade channels the unique advantage of PCM tbat stems from the coding principle is presumably valuable in any transmission medium. We have shown that, theoretically, PCM methods can achieve lower power requirements than any of the other metbods considered and can do so with considerably less frequency space.

While tbis paper is primarily concerned with tbe transmission of multiplex telephony, it seems appropriate to dwell briefly on the transmission of television signals by radio relay. The repeater plan of Fig. 4 is capable of handling long distance transmission of a 5-mc (video) television signal (by FM). The frequency occupancy of a single two-way route is 80 mc. The occupancy for 1000 4-kc telephone channels is 72 mc from Table VI for binary PCM-AM with dual polarization. At this rate a 5-mc video television band would require 90 mc assuming that the 39 db signal-to-quantizing noise ratio is satisfactory for television.[42] Remembering that route congestion can lead to a greater occupancy than 80 mc in the FM case and perhaps to no increase over 90 mc in the PCM case, we conclude tbat on tbese assumptions PCM might be a desirable method for long television relay routes. In the event that a better signal-to-noise ratio is found necessary, binary PCM provides 6 db improvement for each additional digit.

These conclusions relate to the transmission problem under tbe assumed conditions, and do not reflect the impact of many factors that may grow out of an application to a real situation. As bas been said before, this paper should be taken to illustrate tbe way in wbich tbe transmission factors are interrelated, and the philosopby by which the problem is approacbed, rather than to find an unequivocally best system.

In preparing this paper the authors have, of course, drawn on the general transmission background of the Bell Telephone Laboratories. Nourishment has come particularly from W. M. Goodall, A. L. Durkee, H. S. Black, D. H. Ring, J. C. Scbelleng and F. B. Llewellyn in addition to those mentioned specifically in the paper.

We wish specifically to tbank Mr. R. K. Potter, whose broad transmission concepts were responsible for initiating tbe work.

APPENDIX I

Noise in PCM circuits

In the transmission of speech by PCM the kinds of noise and distortion which are acquired by other systems in transmission are completely missing.

[42] W. M. Goodall, "Television by Pulse-Code Modulation." Paper presented at 1949 IRE National Convention, New York, March 9, 1949.

Instead, a special kind of impairment is incurred at the terminals, because of the fact that the speech wave is transmitted by quantized amplitude samples of the wave. Transmitting samples of a wave results in a received wave having no impairment, provided the samples are not subjected to time or amplitude distortion. In PCM, since tbe samples are telegraphed, their reception is inaccurate by the quantization imposed by the code. These errors in the samples constitute the sole inherent impairment in transmission.

Strictly speaking, the transmission impairment in PCM is manifested only when a signal is being transmitted. An imaginary telephone circuit with the transmitting side completely devoid of any kind of signal, except that from the talker, could be transmitted by PCM from coast to coast and would sound completely silent if the talker were silent. In any real situation, however, some background noise (room noise, breath noise or line noise) is always present in tbe subscriber's circuit. This background noise is usually comparable to or greater than the weak parts of weak speech. In order to transmit tbe speech of weak talkers the size of the discrete amplitude steps must be small with the result that at least a few steps are always brought into play by background noise.

Being thus enabled to rule out the case of no *signal* we are able to ascribe a basic signal impairment to a PCM system. This impairment is, strictly speaking, a result of non-linear distortion inherent in the quantizing, but because of its very complex nature it behaves, and sounds, much like thermal noise and we have accordingly called it quantizing *noise*. A PCM circuit can be regarded as a source of noise whose rms value is simply related to the size of the quantizing step and the sampling frequency, as follows:

In a low-pass band extending to approximately 40% of the sampling frequency the basic noise power is related to the power of a sine wave signal by

$$\frac{\text{Signal power}}{\text{Noise power}} = 20 \log_{10} \frac{\text{peak-to-peak signal voltage}}{\text{step voltage}} + 3 \text{ db}$$

This band of noise has an amplitude distribution somewhat different from thermal noise, and a spectral distribution which depends somewhat upon the spectral distribution of the signal and upon its amplitude and disposition with respect to the step boundaries.

For a sine wave signal the noise spectrum is characterized by a number of prominent components rising above a diffuse background of numerous smaller components. The outstanding components may be either harmonics of the signal frequency or differences between harmonics of the sampling frequency and harmonics of the signal frequency. The background thus consists of an array of various orders of cross-products between the signal and the sampling rate. When the amplitude of the signal is comparable to one step in the quantizing process, a few components may contain a substan-

tial portion of the total power in the distortion spectrum. If the signal is not only weak but has its frequency near the low edge of the band, the distortion spectrum has a decided downward slope on the frequency scale with a major part of the distortion power concentrated in the lower harmonics of the signal frequency. Similarly, a weak signal at the upper edge of the band may cause a few scattered difference products to be outstanding. Stronger signals with more centrally located frequencies give practically a uniform distribution of distortion power throughout the signal band. For all except the extreme cases of low amplitudes and frequencies near the edges of the band, the weighting network used to evaluate the telephonic interfering effect of noise gives a reading equal to that obtained with a flat band of thermal noise of the same mean power. The exceptional cases show a spread in the readings which are sensitive to amplitude, frequency and disposition with respect to step boundaries. The spread is reduced when complex signal waves are applied. An operationally significant case is that in which the noise is produced by residual power hum in the equipment. In such a case, weighted noise readings range from approximately the value obtained for flat noise of the same mean power, to several db lower. Connecting even a short subscriber's loop to the input usually adds enough miscellaneous noise, if the steps are as small as they need to be, to remove the variability and to yield a reading within one db of the equivalent flat noise case.

Thus, a PCM system, like any other transmission system, possesses a noise source and experiments show that this noise combines by power addition with that from another system connected to the input or output of the PCM system. In tandem connections of PCM systems in which successive quantizations may occur, the quantizing noise also adds like power, from system to system, and soon becomes almost indistinguishable from thermal noise.

The quantizing noise consists of distortion products which may be classified as two kinds. One class includes those products which would be produced by transmitting the wave through a transducer whose input-output characteristic is stepped like a staircase. If such a transducer were actually used the PCM process would be equivalent to sampling its output at a regular rate and transmitting the step designations by code. This sampling process, applied to the stepped transducer output, produces the other class of distortion (or noise) and is illustrated in Fig. 34. Let us consider the sampled value as the sum of the true value plus the step error, and focus attention on the step error which is responsible for the distortion. At minimum permissible sampling frequency (twice the highest signal frequency), the step errors in consecutive samples are practically unrelated to each other. The low-pass output filter passes most of the power in this sequence of random errors

when they occur at a frequency only twice the filter cutoff frequency. See A in Fig. 34. If the sampling frequency were increased from the minimum permissible value, the consecutive step errors would still be unrelated to each other, and more and more of the step error spectrum (noise) would fall above the low-pass filter. This is shown in B.

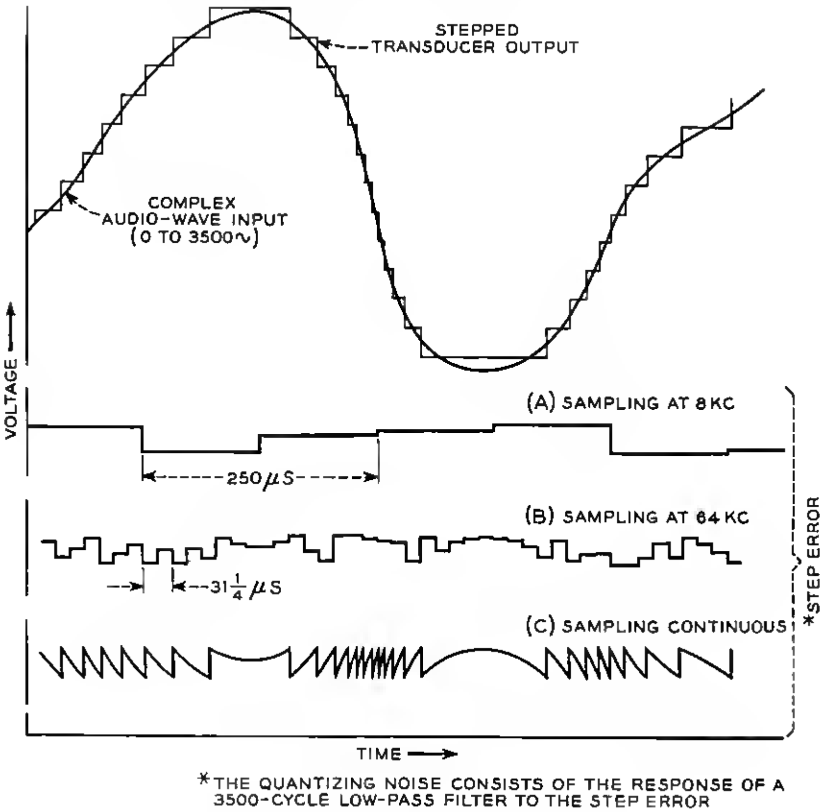


Fig. 34—Stepping and sampling an audio wave.

Reduction of noise would occur in this way until the sampling frequency became so high that a considerable number of samples are taken while the wave crosses a step interval. Correlation between successive step errors then begins to be apparent. When the interval between samples becomes vanishingly small, the process is equivalent to transmitting the stepped transducer output directly. This case appears in C.

In an alternate line of thinking, one may regard the stepped transducer

output as the signal wave plus a wide spectrum of distortion frequencies representing the effect of the steps. From this point of view, it is clear that only a high sampling frequency prevents lower sidebands associated with the sampling frequency and its harmonics from overlapping the signal band.

Quantizing noise decreases with increase of sampling frequency at an initial rate of approximately 3 db per octave and continues until correlation of successive errors becomes appreciable. This occurs at a sampling frequency which is dependent upon the spectral distribution of the signal, being lower for signals having a predominately low-frequency spectral density. An increase of step size also reduces the lowest sampling frequency at which effects of correlation are observed. Figure 35 shows curves calculated for an

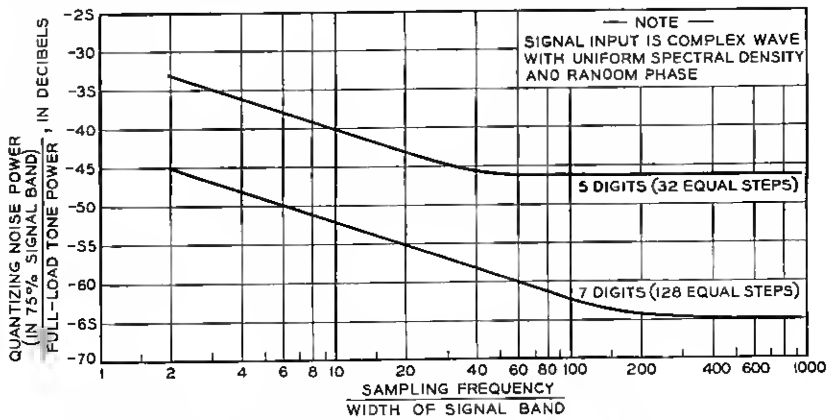


Fig. 35—Variation of quantizing noise with sampling frequency.

input consisting, in fact, of thermal noise. Such an input is a rough approximation to a speech wave.

The asymptotic values shown for five and seven digits represent the quantizing noise corresponding to transmission of the thermal noise signal through stepped transducers having 32 and 128 steps, respectively. The curves suggest that sampling is a penalty such that 32-step granularity without sampling is about equivalent[43] to 128-step granularity with sampling at the minimum rate. However, sending information which designates the irregular instants of time at which the signal enters and leaves each step interval is far less efficient than designating the steps at the regular instants of the minimum sampling rate.

[43] The equivalence would be in terms of total noise power; the properties of the asymptotic noise are different than were described earlier in this appendix, for sampling at the minimum rate.

APPENDIX II

Interference between two frequency modulated waves

This problem occurs so frequently in the present paper that its solution is appended here for reference. Figure 36 shows a geometric figure from which the phase of a two-component wave can be calculated. We write

$$P \cos \theta + Q \cos \varphi = R \cos \psi$$

where

$$R^2 = P^2 + Q^2 + 2PQ \cos (\theta - \varphi)$$

$$\tan \psi = \frac{P \sin \theta + Q \sin \varphi}{P \cos \theta + Q \cos \varphi}$$

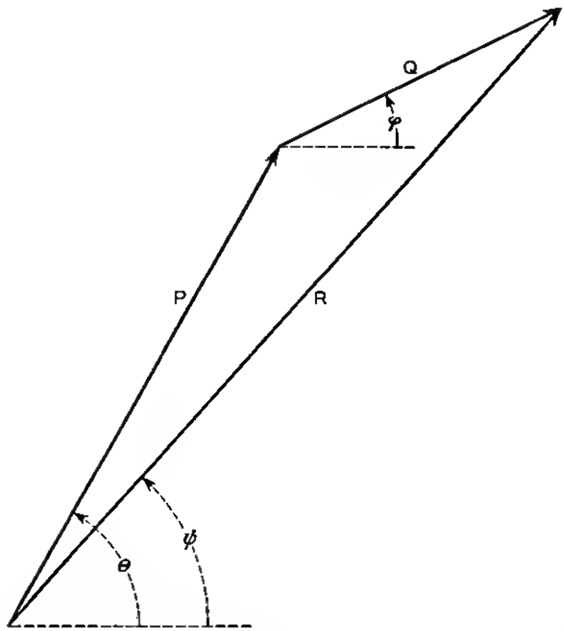


Fig36. —Geometric solution for resultant phase of two frequency modulated waves.

The response of a perfect frequency detector in radians/sec. is given by

$$\Omega = d\psi/dt = \frac{d}{dt}\left(\text{arc tan} \frac{P \sin \theta + Q \sin \varphi}{P \cos \theta + Q \cos \varphi}\right)$$

$$= \tfrac{1}{2}(\theta' + \varphi') + \frac{\theta' - \varphi'}{2} \frac{P^2 - Q^2}{P^2 + Q^2 + 2PQ \cos (\theta - \varphi)}$$

In the above expression, the primes represent derivatives with respect to time. If $Q/P < 1$, we expand in Fourier series, obtaining

$$\Omega = \theta' + (\theta' - \varphi') \sum_{m=1}^{\infty} \left(-\frac{Q}{P}\right)^m \cos m(\theta - \varphi).$$

When Q/P is small, we retain only the term proportional to Q/P as the error, which may be written in the compact form:

$$\Omega - \theta' \doteq \frac{Q}{P} \frac{d}{dt} \sin (\varphi - \theta)$$

If the waves are unmodulated, $\theta = pt$ and $\varphi = qt$, giving

$$\Omega - \theta' \doteq \frac{Q}{P} (q - p) \cos (q - p)t$$

APPENDIX III

PCM FOR BANDWIDTH REDUCTION

We have treated PCM as a means of increasing bandwidth beyond the value corresponding to one pulse per sample per channel (quantized PAM) and have studied the transmission advantages that accrue therefrom. The PCM method can, in principle, serve to reduce bandwidth. An example of bandwidth reduction,[44, 45] suggested to the writers by C. E. Shannon, is as follows:

Any number, say N, of 4 kc telephone channels can be transmitted in the form of one quantized pulse per 125 microseconds, by sampling all channels in the usual way, encoding each sample into a code symbol having, say, 64 possible values, assembling all code pulses into one new group and decoding this group at the transmitter. If only one channel were to be transmitted the decoded signal would have 64 possible amplitudes; for two channels it would have 64^2 possible amplitudes, and for N channels, 64^N. Now, if a single quantized pulse conveying these amplitudes could be transmitted without an error as large as one step, the receiver could encode the quantized pulse, disassemble the resulting code pulses into groups according to channels and decode the groups to obtain the N channel samples. The requirements on transmission circuits capable of the precision required to transmit even two channels in place of one are very severe, however.

In the event the signals to be transmitted were not speech signals but a very elemental kind of signal such as a black and white pattern requiring for

[44] A paper "Reducing Transmission Bandwidth" by Bailey and Singleton *Electronics*, Aug. 1948 gives a somewhat different example of reduction.

[45] An early disclosure of a system theoretically capable of any desired amount of bandwidth reduction is contained in U. S. Patent No. 2,056,284, Oct. 6, 1936, issued to L. A. MacColl. As in the current proposals, the decreased band is obtained at the expense of a vastly greater signal-to-noise ratio requirement and the necessity for precise synchronism between transmitter and receiver.

its description not 64 values but only 2, the number of possible amplitudes would be 2^N. With some of the better transmission circuits in existence, as many as 10 such channels could be multiplexed in the same bandwidth required by one cbannel, by this adaptation of the PCM method.

The above considerations show that PCM offers the means of matching the transmission signal to the capabilities of the transmission circuit in order to transmit tbe maximum amount of information. As has been shown, with microwave telephone systems the economical balance seems to come well over on the wide band side, permitting operation with low transmitted power through relatively strong interference.

APPENDIX IV

Supplementary details of derivation of bandwidth-curves

PPM-AM

The diagram of Fig. 8 shows that the maximum time deviation is assumed to be

$$\epsilon = \frac{1}{2Nf_r} - T = \frac{1}{2Nf_r} - \frac{1}{F_b}. \tag{1}$$

The shift in time produced by an interfering voltage of magnitude E_n at the slicing instant is

$$\Delta t = E_n/s \tag{2}$$

where s is the slope of the signal pulse at the slicing instant. For small noise the slicing instant occurs near half the peak, E, of the pulse and the slope of the assumed sinusoidal pulse (Fig. 6) is:

$$s = \frac{\pi}{2T} E = \frac{\pi F_b E}{2} \tag{3}$$

Hence

$$\Delta t = \frac{2E_n}{\pi F_b E}. \tag{4}$$

The signal-to-noise power ratio in the channel is the ratio of mean square values of signal deviation σ and Δt. For thermal noise we assume that the root mean square value is one fourth the peak and place tbe peak at $1/\sqrt{2}$ times $E/2$ for marginal operation. We write, therefore,

$$E_n = E/2\sqrt{2}, \qquad \overline{\Delta t^2} = (4E_n^2/\pi^2 F_b^2 E^2)/16 \tag{5}$$

$$\overline{\sigma^2} = \frac{\epsilon^2}{2} = \frac{1}{2}\left(\frac{1}{2Nf_r} - \frac{1}{F_b}\right)^2 = \frac{1}{2F_b^2}\left(\frac{F_b}{2Nf_r} - 1\right)^2 \tag{6}$$

$$S/N = 4\pi\left(\frac{F_b}{2Nf_r} - 1\right) = 4\pi\left(\frac{B}{4Nf_r} - 1\right). \tag{7}$$

Here S/N is the ratio of root mean square audio signal and noise voltages. The formula illustrates a general principle common to all the pulse systems in that the marginal signal-to-noise ratio is a function of B/Nf_r. The axes of the curves have been labeled for $N = 1000$ and $f_r = 8$ kc, but can be read for any other values of N and f_r by changing B accordingly. Equation (7) was used to plot the marginal power curve of Fig. 9. We note that tbe ratio of rms pulse voltage at the top of the pulse to the rms noise voltage is $E/\sqrt{2}$ divided by $E_n/4$ which leads to a value of eight when (5) is substituted. The "slicer advantage" is thus the right-hand member of (7) divided by eight.

For CW interference in a PPM-AM system the procedure used above applies except that the root mean square interference is now $1/\sqrt{2}$ times the peak instead of one fourth. The marginal ratio of rms audio signal to rms audio interference ratio is therefore poorer by a factor of $2\sqrt{2}$, or

$$S/I = \pi\sqrt{2}\left(\frac{B}{4Nf_r} - 1\right). \tag{8}$$

When the interference is from a similar system, we calculate the distribution of the disturbance as follows. The probability that there is an interfering pulse present during slicing is the ratio of the pulse duration to the channel allotment, or

$$p_0 = 2TNf_r = \frac{2Nf_r}{F_b}. \tag{9}$$

The interfering carrier will not, in general, he exactly synchronous with the wanted carrier, and hence the actual interference is a heat frequency with envelope having a voltage distribution calculable from the pulse shape. For a sinusoidal pulse of height A, the probahility $p(y)\,dy$ that the instantaneous magnitude of the interfering envelope is in the interval dy at y is

$$p(y)\,dy = \frac{p_0\,dy}{\pi\sqrt{y(A - y)}}, \qquad 0 < y < A. \tag{10}$$

Since the relative phase of the two carriers drifts with time, the mean square interfering voltage is half the mean square interfering envelope, or

$$E_n^2 = \tfrac{1}{2}\int_0^A y^2 p(y)\,dy = \frac{3Nf_rA^2}{9F_b} = \frac{3Nf_rA^2}{4B}. \tag{11}$$

Hence

$$\overline{\Delta t^2} = \frac{12Nf_rA^2}{\pi^2 B^3 E^2} \tag{12}$$

and

$$\frac{S}{I} = \frac{\pi E}{A}\left(\frac{B}{4Nf_r} - 1\right)\sqrt{\frac{B}{6Nf_r}}. \tag{13}$$

For marginal interference $E = 2\sqrt{2}\,A$ and

$$\frac{S}{I} = 2\pi\left(\frac{B}{4Nf_r} - 1\right)\sqrt{\frac{B}{3Nf_r}}. \tag{14}$$

This equation shows that S/I varies as $(B/Nf_r)^{3/2}$ for large bandwidtbs giving 9 db improvement per octave of bandwidth. The curves of Fig. 10 were plotted from equations (8) and (14).

PPM-FM

The pulse here is transmitted by a change in frequency from f_1 to $f_1 + \beta$ and back again. The total frequency swing β corresponds to the pulse height E in the AM case. The frequency detector delivers a pulse of height β to the baseband filter. Associated with the pulse is the error caused by noise or interference in the rf-band. In the case of fluctuation noise having mean power P_n per cycle in the rf-medium, a baseband filter of width F_b accepts the familiar triangular voltage distribution of noise with frequency resulting[46] in a mean square integrated magnitude expressed on a frequency scale as:

$$\overline{E_n^2} = P_n F_b^3/3W_c \tag{15}$$

where W_c is the mean carrier power. Then, on substituting β for E, and the above expression for $\overline{E_n^2}$ in the equation for Δt:

$$\overline{\Delta t^2} = \frac{4P_n F_b}{3\pi^2\beta^2 W_c} \tag{16}$$

Taking the ratio of $\overline{\epsilon^2}$ to $\overline{\Delta t^2}$,

$$(S/N)^2 = \frac{3\pi^2\beta^2 W_c}{8P_n F_b^3}\left(\frac{F_b}{2Nf_r} - 1\right)^2. \tag{17}$$

The radio signal bandwidth B is approximately equal to the frequency swing plus a sideband at each end or

$$B = \beta + 2F_b \tag{18}$$

Using this relation to eliminate β, we have

$$(S/N)^2 = \frac{3\pi^2(B - 2F_b)^2 W_c}{8P_n F_b^3}\left(\frac{F_b}{2Nf_r} - 1\right)^2. \tag{19}$$

For marginal operation of the FM limiter:

$$W_c = kP_nB \tag{20}$$

wbere we shall assume $k = 16$ in numerical calculations.

[46] An elementary component of interference $Q\cos qt$ produces a frequency error $(Q/P)f\cos 2\pi ft$ where f is the difference between the interfering and carrier frequencies. The corresponding mean square frequency error is $f^2Q^2/2P^2$. But $Q^2/2 = P_n df$ and there are equal contributions from upper and lower sidebands centered around the carrier. Also replacing $P^2/2$ by W_c, we get a mean square frequency error in band df at f equal to $P_n f^2 df/W_c$. Integrating over frequencies from 0 to F_b gives the above result.

Then

$$(S/N)^2 = \frac{3k\pi^2}{8}\left(\frac{B}{F_b}\right)^3\left(1 - 2\frac{F_b}{B}\right)^2\left(\frac{F_b}{2Nf_r} - 1\right)^2. \tag{21}$$

The signal-to-noise ratio is found to be maximum when

$$\frac{F_b}{Nf_r} = \sqrt{\left(\frac{B}{4Nf_r} + 1\right)^2 + \frac{3B}{Nf_r}} - \frac{B}{4Nf_r} - 1. \tag{22}$$

To calculate the CW interference with an idle channel we assume that the carrier wave of the system is represented by

$$V_1(t) = P\cos\left[2\pi f_1 t + \phi(t)\right] \tag{23}$$

where

$$\phi(t) = \begin{cases} \pi\beta\left(t + \dfrac{1}{\pi F_b}\sin\pi F_b t\right), & 0 < t < \dfrac{1}{F_b} \\ \pi\beta/F_b, & \dfrac{1}{F_b} < t < \dfrac{1}{2Nf_r} \end{cases} \tag{24}$$

$$\phi(-t) = -\phi(t), \qquad \phi\left(t \pm \frac{2m}{F_b}\right) = \phi(t), \qquad m = 1, 2, \cdots. \tag{25}$$

We have chosen $\phi(t)$ so that the phase is a continuous function of time with a derivative representing the correct frequency. This gives a sinusoidal change in the instantaneous frequency $\phi'(t)/2\pi$ starting with the value f_1 at $t = -\frac{1}{F_b}$, reaching the peak $f_1 + \beta$ at $t = 0$, and subsiding to f_1 at $t = \frac{1}{F_b}$. By making the wave repeat with frequency Nf_r, we assume all channels of the system are idle since all pulses are at their central positions. It seems reasonable to neglect the effect of variations in adjacent channel loading on CW interference in one channel. The interfering CW wave is represented by

$$V_2(t) = Q\cos(2\pi f_2 t + \theta) \tag{26}$$

To a first approximation the resulting error in frequency at the output of the frequency detector is:

$$\delta(t) = \frac{Q}{2\pi P}\frac{d}{dt}\sin\left[2\pi(f_1 - f_2)t + \phi(t) - \theta\right] \tag{27}$$

By straghtforward Fourier series expansion and differentiation:

$$\delta(t) = \frac{Q}{P}F_b\sum_{n=-\infty}^{\infty}(c + n\lambda)A_n\cos\left[2\pi F(c + n\lambda)t - \theta\right] \tag{28}$$

where:

$$c = (f_1 - f_2)/F_b, \qquad \lambda = Nf_r/F_b, \qquad y = \beta/F_b, \tag{29}$$

$$A_n = 2\lambda \mathcal{J}_{2n\lambda - y}(y) - \frac{1}{n\pi}[(-)^n + \sin(2n\lambda - y)\pi], \tag{30}$$

$$A_0 = 2\lambda \mathcal{J}_{-y}(y) + (1 - 2\lambda)\cos \pi y. \tag{31}$$

The function $\mathcal{J}_\nu(y)$ is Anger's function:[47]

$$\mathcal{J}_\nu(y) = \frac{1}{\pi}\int_0^\pi \cos(\nu\theta - y\sin\theta)\,d\theta \tag{32}$$

It is equal to $J_\nu(y)$, the more familiar Bessel function of the first kind, only when ν is an integer. The values of $\nu = 2n\lambda - y$ appearing in this solution are in general not integers and hence the ordinary tables of Bessel functions are inapplicable.

The baseband filter accepts the components of the error which have frequencies in the range:

$$-F_b \leq F_b\,(c + n\lambda) \leq F_b \tag{33}$$

or

$$-\frac{1 + c}{\lambda} \leq n \leq \frac{1 - c}{\lambda}. \tag{34}$$

The interfering wave in the baseband filter output is then

$$\delta_0(t) = \frac{Q}{P}F_b \sum_{n=n_1}^{n_2} (c + n\lambda)A_n \cos[2\pi F_b(c + n\lambda)t - \theta] \tag{35}$$

where n_1 is the smallest integer not less than $-(1 + c)/\lambda$ and n_2 is the largest integer not greater than $(1 - c)/\lambda$. It would be convenient at this point to assume that $\overline{\Delta t^2}$ is expressible directly in terms of $\overline{\delta_0^2(t)}$. However, there is reason to believe that such an assumption is pessimistic especially at the higher bandwidths where the disturbance $\delta_0(t)$ may never reach its maximum values in the neighborhood of the actual slicing instant. A complete investigation requires a study of the instantaneous wave form of $\delta_0(t)$ in the neighborhood of the slicing instant. We note that if the slicer operates at the trailing edge, the unperturbed slicing instant is $t = \dfrac{1}{2F_b} + m/f_r$, and the value of the disturbance at that instant is:

$$\delta_0\left(\frac{1}{2F_b} + m/f_r\right) = \frac{Q}{P}F_b \sum_{n=n_1}^{n_2} (c + n\lambda)A_n$$

[47] Watson, Theory of Bessel Functions, Chapter X.

$$\cos\left[2\pi(c + n\lambda)\left(\frac{1}{2} + m\frac{F_b}{f_r}\right) - \theta\right], \qquad m = 0, \pm 1, \pm 2, \cdots \tag{36}$$

Averaging the square over all values of $\left[2m\pi\frac{F_b}{f_r} - \theta + c\pi\right]$, which may be treated as a randomly distributed angle, we find an expression for δ_0^2 averaged over the values at the slicing instants and not over all time, viz.:

$$\overline{\delta_0^2} = \frac{Q^2}{2P^2}F_b^2(R^2 + X^2) \tag{37}$$

where

$$R = \sum_{n=n_1}^{\infty}(c + n\lambda)A_n\cos\left[n\pi\lambda\left(1 + \frac{2mF_b}{f_r}\right)\right] \tag{38}$$

$$X = \sum_{n=n_1}^{\infty}(c + n\lambda)A_n\sin\left[n\pi\lambda\left(1 + \frac{2mF_b}{f_r}\right)\right] \tag{39}$$

Then

$$\overline{\Delta t^2} = \frac{2Q^2(R^2 + X^2)}{\pi^2P^2\beta^2} \tag{40}$$

and

$$\frac{S}{I} = \frac{\pi P}{Q}\left(\frac{B}{2F_b} - 1\right)\left(\frac{F_r}{2Nf_r} - 1\right)(R^2 + X^2)^{-1/2} \tag{41}$$

For each value of B, the value of S/I should be computed over a range of values of c; and the lowest value of S/I, corresponding to the most unfavorable allocation of the CW frequency, plotted as a point on the curve. The curve of Fig. 12 was not calculated in this way but estimated from the simpler solution existing when the pulses are contiguous.

When the interference is from a similar system, we substitute for the interfering wave:

$$V_2(t) = Q\cos[2\pi f_1t + \phi(t - \tau) + \theta] \tag{42}$$

This gives

$$\delta(t) = \frac{Q}{2\pi P}\frac{d}{dt}\sin[\phi(t) - \phi(t - \tau) - \theta + 2\pi(f_1 - f_2)t] \tag{43}$$

We distinguish between the cases of overlapping and non-overlapping pulses.

If the pulses do not overlap, we take the origin of time midway between pulses and write

$$\frac{\phi(t) - \phi(t - \tau)}{\pi\beta} = \left\{ \begin{array}{c} -\frac{1}{F_b}, \; -\frac{1}{2Nf_r} < t < -\frac{\tau}{2} - \frac{1}{F_b} \\ t + \frac{\tau}{2} + \frac{1}{\pi F_b} \sin \pi F_b \left(t + \frac{\tau}{2} \right), \; -\frac{1}{F_b} < t + \frac{\tau}{2} < \frac{1}{F_b} \\ \frac{1}{F_b}, \; \frac{1}{F_b} - \frac{\tau}{2} < t < \frac{\tau}{2} - \frac{1}{F_b} \\ -t + \frac{\tau}{2} - \frac{1}{\pi F_b} \sin \pi F_b \left(t - \frac{\tau}{2} \right), \; -\frac{1}{F_b} < t - \frac{\tau}{2} < \frac{1}{F_b} \\ -\frac{1}{F_b}, \; \frac{1}{F_b} + \frac{\tau}{2} < t < \frac{1}{2Nf_r} \end{array} \right\} \tag{44}$$

If the pulses overlap, we also take the origin midway between pulses, but we then obtain

$$\frac{\phi(t) - \phi(t - \tau)}{\pi\beta} = \left\{ \begin{array}{c} -\frac{1}{F_b}, \; -\frac{1}{2Nf_r} < t < -\frac{\tau}{2} - \frac{1}{F_b} \\ t + \frac{\tau}{2} + \frac{1}{\pi F_b} \sin \pi F_b \left(t + \frac{\tau}{2} \right), \; -\frac{1}{F_b} - \frac{\tau}{2} < t < \frac{\tau}{2} - \frac{1}{F_b} \\ t - \frac{1}{F_b} + \frac{2}{\pi F_b} \sin \pi \frac{F_b \tau}{2} \cos \pi F_b t, \; \frac{\tau}{2} - \frac{1}{F_b} < t < \frac{1}{F_b} - \frac{\tau}{2} \\ -t + \frac{\tau}{2} - \frac{1}{\pi F_b} \sin \pi F_b \left(t - \frac{\tau}{2} \right), \; \frac{1}{F_b} - \frac{\tau}{2} < t < \frac{1}{F_b} + \frac{\tau}{2} \\ -\frac{1}{F_b}, \; \frac{1}{F_b} + \frac{\tau}{2} < t < \frac{1}{2Nf_r} \end{array} \right\} \tag{45}$$

In both cases the right-hand members are even functions of t.
Hence

$$\sin\,[\phi(t) - \phi(t - \tau) - \theta] = \sum_{m=1}^{\infty} B_m \cos 2m\pi N f_r t \tag{46}$$

$$\cos\left[\phi(t) - \phi(t - \tau) - \theta\right] = \frac{A_0}{2} + \sum_{m=1}^{\infty} A_m \cos 2m\pi N f_r t \qquad (47)$$

The coefficients A_m and B_m are considerably more complicated than in the corresponding CW case.

PAM-FM

An idle N-channel system sends sinusoidal pulses of duration $2/F_b = 1/Nf_r$ which merge into a continuous sinusoidal variation of frequency expressible by

$$f = f_1 + \frac{\beta}{4} \cos \pi F_b t \qquad (48)$$

where β is the peak-to-peak frequency swing and f_1 is the mid-frequency. A full load audio tone frequency $\frac{q}{2\pi}$ impressed on one cbannel produces a series of sinuosidal pulses of varying heights expressed with sufficient accuracy for a large number of channels by

$$F = f_1 + \frac{\beta}{2} g_0(t) \cos qt \qquad (49)$$

where $g_0(t)$ represents a pulse of unit height and duration $\frac{2}{F_b}$ repeated periodically at the frame frequency f_r. Aperture effect is neglected in this approximation. We may expand $g_0(t)$ in a Fourier series:

$$g_0(t) = \frac{G_0}{2} + \sum_{m=1}^{\infty} G_m \cos 2m\pi f_r t \qquad (50)$$

where

$$G_m = 2f_r \int_{-1/2Nf_r}^{1/2Nf_r} g_0(t) \cos 2m\pi f_r t \, dt \qquad (51)$$

When a rectangular gate of full-channel duration $1/Nf_r$ is used between the baseband filter and the audio output filter for the channel, F represents the input to tbe channel filter. The only term passed by the latter is

$$F_q = \frac{\beta}{4} G_0 \cos qt \qquad (52)$$

$$G_0 = 2f_r \int_{-1/2Nf_r}^{1/2Nf_r} \tfrac{1}{2}(1 + \cos 2\pi N f_r t) \, dt = \frac{1}{N} \qquad (53)$$

Therefore the peak sine wave channel output is $\beta/4N$, and the mean square value is $\beta^2/32N^2$. If a gating function $g_1(t)$ is used instead of a rectangular gate, we replace $g_0(t)$ by $g_0(t)g_1(t)$ in the calculation of G_0.

When the interference is fluctuation noise of mean power $P_n df$ in bandwidth df, the mean square frequency error in the output of the frequency detector is $w_n(f)df$ at frequency f, where

$$w_n(f) = P_n f^2/W_c \tag{54}$$

The baseband filter accepts the portion of this spectrum between $f = 0$ and $f = F_b$.

The action of the rectangular gate on this spectrum may be calculated by multiplication of the typical spectral component by the gating function:

$$G(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos 2\pi m f_r t \tag{55}$$

where for $G(t) = 1$ throughout the channel allotment time,

$$a_m = 2f_r \int_{-1/2Nf_r}^{1/2Nf_r} \cos 2m\pi f_r t \, dt = \frac{2 \sin \dfrac{m\pi}{N}}{m\pi} \tag{56}$$

Each harmonic of $G(t)$ beats with the noise spectrum on either side to produce audio components which sum up to total mean square audio noise:

$$W_n = \frac{1}{4} \int_0^{f_r/2} \left[a_0^2 w_n(f) + \sum_{m=1}^{2N} a_m^2 w_n(mf_r + f) + \sum_{m=1}^{2N} a_m^2 w_n(mf_r - f) \right] df \tag{57}$$

The summations stop at $2N$ because the baseband filter cuts off at $f = F_b = 2Nf_r$. The contribution of the a_0 term is negligible. Then

$$W_n = \frac{P_n}{\pi^2 W_c} \sum_{m=1}^{2N} \frac{\sin^2 \dfrac{m\pi}{N}}{m^2} \int_0^{f_r/2} [(mf_r + f)^2 + (mf_r - f)^2] \, df = \frac{P_n f_r^3}{\pi^2 W_c} \sum_{m=1}^{2N} \left(1 + \frac{1}{12m^2}\right) \sin^2 \frac{m\pi}{N}. \tag{58}$$

When N is large, the sum approaches

$$W_n \cong \frac{N P_n f_r^3}{\pi^2 W_c} \tag{59}$$

and

$$(S/N)^2 = \frac{\pi^2 W_c}{32 N^3 P_n f_r} \left(\frac{\beta}{f_r}\right)^2 \tag{60}$$

$$S/N = \frac{\pi}{4} \left(\frac{B}{Nf_r} - 4\right) \sqrt{\frac{kB}{2Nf_r}} \tag{61}$$

on substituting $W_c = kP_nB$, $\beta = B - 2F_b$, and $F_b = 2Nf_r$. Equation (61) with $k = 16$ was used to obtain the marginal power curve of Fig. 13. The result may be compared with that given by Rauch[48] (See also Landon[4]) for a rectangular pulse and rectangular gate, which requires a higher value for F_b. The two systems give the same signal-to-noise ratio when the rectangular pulse and gate of Rauch's system endure for one half of the total channel allotment time. The value of F_b necessary for such a case was estimated by Rauch to be $3.5Nf_r$. A calculation made as above, except that the gate was assumed sinuosoidal instead of rectangular, showed very nearly the same signal-to-noise ratio.

The case of CW interference with all channels idle is represented by the r.f. wave:

$$V(t) = P \cos\left(2\pi f_1 t + \frac{\beta}{2F_b} \sin \pi F_b t\right) + Q \cos 2\pi f_2 t. \tag{62}$$

When Q/P is small, the detected frequency is

$$\begin{aligned} f &= f_1 + \frac{\beta}{4} \cos \pi F_b t - \frac{Q}{2\pi P} \frac{d}{dt} \sin\left[2\pi(f_1 - f_2)t + \frac{\beta}{2F_b} \sin \pi F_b t\right] \\ &= f_1 + \frac{\beta}{4} \cos \pi F_b t - \delta(t). \end{aligned} \tag{63}$$

By Fourier series expansion followed by differentiation, the error $\delta(t)$ may be written as:

$$\delta(t) = \frac{Q}{P} \sum_{m=-\infty}^{\infty} \left(f_1 - f_2 + \frac{mF_b}{2}\right) J_m(x) \cos 2\pi \left(f_1 - f_2 + \frac{mF_b}{2}\right) t, \tag{64}$$

where $x = \beta/2F_b$. The baseband filter passes only those components of $\delta(t)$ which have frequencies in the range $-F_b$ to F_b. Writing $c = (f_1 - f_2)\, F_b$, we find:

$$\delta_0(t) = \frac{QF_b}{P} \sum_{m=m_1}^{m_2} \left(c + \frac{m}{2}\right) J_m(x) \cos 2\pi F_b \left(c + \frac{m}{2}\right) t \tag{65}$$

where m_2 is the largest integer which does not exceed $2(1 - c)$ and m_1 is the smallest integer which is not exceeded by $-2(1 + c)$. The term integer is here understood to include zero and both positive and negative integers. The wave $\delta_0(t)$ is next multiplied by the gating function $G(t)$ and the components falling in the audio band $-f_r/2$ to $f_r/2$ selected. We find:

$$G(t)\delta_0(t) = \frac{F_b Q}{2P} \sum_{m=m_1}^{m_2} \sum_{n=-\infty}^{\infty} a_n \left(c + \frac{m}{2}\right) J_m(x) \cos 2\pi[(2c + m)N + n] f_r t \tag{66}$$

with $a_n = a_{-n}$.

[4] Loc. cit.
[48] L. L. Rauch, Fluctuation Noise in Pulse-Height Multiplex Radio Links, Proc. I.R.E., Vol. 35, Nov. 1947, pp. 1192–1197.

For each value of m, there is only one value of n satisfying the audio filter inequality, which may be written:

$$-\tfrac{1}{2} - (2c + m)N < n < \tfrac{1}{2} - (2c + m)N \tag{67}$$

Hence interference accepted by the channel filter is:

$$I(t) = \frac{F_b Q}{2P} \sum_{m=m_1}^{m_2} a_n \left(c + \frac{m}{2}\right) J_m(x) \cos 2\pi[(2c + m)N + n]f_r t \tag{68}$$

The mean square value of interference is

$$\overline{I^2(t)} = \frac{F_b^2 Q^2}{8P^2} \sum_{m=m_1}^{m_2} a_n^2 \left(c + \frac{m}{2}\right)^2 J_m^2(x). \tag{69}$$

The signal-to-interference ratio is

$$(S/I)^2 = \frac{G_0^2 \beta}{32} [\overline{I^2(t)}]^{-1} \tag{70}$$

or

$$S/I = \frac{G_0 \beta P}{2F_b Q} \left[\sum_{m=m_1}^{m_2} a_n^2 \left(c + \frac{m}{2}\right)^2 J_m^2(x)\right]^{-1/2}. \tag{71}$$

When a rectangular gate of duration equal to the full channel allotment is used, we substitute $a_n = 2(\sin n\pi/N)/n\pi$. We then find that the largest values of mean square interference occur when c is an odd multiple of one fourth. If we set

$$c = -(2r + 1)/4, r = 0, \pm 1, \pm 2, \cdots \tag{72}$$

it follows that if N is an even integer,

$$n = (r + \tfrac{1}{2} - m)N, \tag{73}$$

$$\sin \frac{n\pi}{N} = \sin (r + \tfrac{1}{2} - m)\pi, \tag{74}$$

$$\left|\sin \frac{n\pi}{N}\right| = 1. \tag{75}$$

Substituting these values in the expression for S/I, we find

$$S/I = \frac{\pi \beta P}{2F_b Q} \left[\sum_{m=r-1}^{r+2} J_m^2(x)\right]^{-1/2}. \tag{76}$$

The value of r is to be chosen as the integer which makes S/I a minimum; i.e., we place the CW frequency at that part of the band where it does the most damage. The curve marked CW(Gate) of Fig. 14 was obtained in this way.

When instantaneous sampling is used instead of a gate, the value of a_n becomes a constant for all values of n of interest and is equal to a_0. We then find:

$$S/I = \frac{\beta P}{2F_b Q}\left[\sum_{m \geq -2(1+c)}^{\leq 2(1-c)} \left(c + \frac{m}{2}\right)^2 J_m^2(x)\right]^{-1/2}. \tag{77}$$

Here c is to be selected to make S/I minimum for each value of x. The poorest values of S/I occur when Bessel functions of comparable order and argument appear in the summation, which means that c is in the neighborhood of $-x/2$. The corresponding difference between the CW and mid-carrier frequencies is one-fourth the peak-to-peak swing.

To calculate the interference between two similar idle systems, we set

$$\begin{aligned} V(t) &= P \cos\left(2\pi f_1 t + \frac{\beta}{2F_b} \sin \pi F_b t\right) \\ &\quad + Q \cos\left[2\pi f_2 t + \frac{\beta}{2F_b} \sin (\pi F_b t - \theta)\right]. \end{aligned} \tag{78}$$

The frequency error registered in the first system is then

$$\begin{aligned} \delta(t) &= \frac{Q}{2\pi P}\frac{d}{dt} \sin [2\pi(f_1 - f_2)t + x \sin \pi F_b t - x \sin (\pi F_b t - \theta)] \\ &= \frac{Q}{2\pi P}\frac{d}{dt} \sin \left[2\pi(f_1 - f_2)t + 2x \sin \frac{\theta}{2} \cos\left(\pi F_b t - \frac{\theta}{2}\right)\right] \\ &= \frac{Q}{P}\sum_{m=-\infty}^{\infty}\left(f_1 - f_2 + \frac{mF_b}{2}\right) J_m\left(2x \sin \frac{\theta}{2}\right) \\ &\qquad \cdot \cos\left[2\pi\left(f_1 - f_2 + \frac{mF_b}{2}\right)t - \frac{m\theta}{2}\right]. \end{aligned} \tag{79}$$

It follows that the response of the baseband filter is

$$\delta_0(t) = \frac{QF_b}{P}\sum_{m=m_1}^{m_2}\left(c + \frac{m}{2}\right) J_m\left(2x \sin \frac{\theta}{2}\right) \cos\left[2\pi F_b\left(c + \frac{m}{2}\right)t - m\frac{\theta}{2}\right]. \tag{80}$$

The effects of the channel gate and filter are computed in the same way as for CW, giving the audio interference:

$$\begin{aligned} I(t) &= \frac{F_b Q}{2P}\sum_{m=m_1}^{m_2} a_n \left(c + \frac{m}{2}\right) J_m\left(2x \sin \frac{\theta}{2}\right) \\ &\qquad \cdot \cos\left(2\pi[(2c + m)N + n]f_r t - m\frac{\theta}{2}\right) \end{aligned} \tag{81}$$

Since the two systems occupy the same frequency assignment, we assume that c is not greatly different from zero. Then for fixed θ:

$$\overline{I^2(t)} = \frac{F_b^2 Q^2}{32P^2} \sum_{m=m_1}^{m_2} m^2 a_n^2 J_m^2 \left(2x \sin \frac{\theta}{2}\right) \tag{82}$$

Since θ, the frame phase difference, is a random angle we average over its possible values by setting:

$$\begin{aligned} A_m(x) &= \frac{1}{2\pi} \int_0^{2\pi} J_m^2 \left(2x \sin \frac{\theta}{2}\right) d\theta = \frac{1}{\pi} \int_0^{\pi} J_m^2(2x \sin \phi)\, d\phi \\ &= \frac{\Gamma^2(m + \frac{1}{2})(2x)^{2m}}{\pi(2m)!(m!)^2} \\ &\quad \cdot {}_2F_3(m + \tfrac{1}{2}, m + \tfrac{1}{2}; 2m + 1, m + 1, m + 1; -4x^2), \\ &\qquad\qquad m \geq 0 \end{aligned} \tag{83}$$

$$A_{-m}(x) = A_m(x)$$

Noting further that for $c = 0$, $m_1 = -2$, $m_2 = 2$, and $n = -mN$, we have then:

$$\overline{I^2(t)} = \frac{F_b^2 Q^2}{16P^2} [a_N^2 A_1(x) + 4a_{2N}^2 A_2(x)] \tag{84}$$

and

$$S/I = \frac{G_0 \beta P}{F_b Q \sqrt{2}} [a_N^2 A_1(x) + 4a_{2N}^2 A_2(x)]^{-1/2}. \tag{85}$$

For a sinusoidal pulse and rectangular gate of full channel allotment time in duration, $a_N = a_{2N} = 0$, $I^2(t) = 0$, and S/I is infinite. If instantaneous sampling is used,

$$G_0 = a_N = a_{2N} \tag{86}$$

and

$$S/I = \frac{\beta P}{\sqrt{2}\, F_b Q} [A_1(x) + 4A_2(x)]^{-1/2} \tag{87}$$

The curve for similar system interference with instantaneous sampling, Fig. 14, was calculated from Eq. (87).

For small values of x, we may use the ascending power series:

$$A_1(x) = \frac{x^2}{2} \left[1 - \frac{3^2}{3 \cdot 2^2 \cdot 1!} x^2 + \frac{3^2 \cdot 5^2}{3 \cdot 4 \cdot 2^2 \cdot 3^2 \cdot 2!} x^4 - \cdots \right] \tag{88}$$

$$A_2(x) = \frac{3x^4}{32} \left[1 - \frac{5^2}{5 \cdot 3^2 \cdot 1!} x^2 + \frac{5^2 \cdot 7^2}{5 \cdot 6 \cdot 3^2 \cdot 4^2 \cdot 2!} x^4 - \cdots \right] \tag{89}$$

For large values of x, the following asymptotic formula has been derived by Mr. S. O. Rice by use of the Mellin-Barnes contour integral representation and the method of steepest descents:

$$\pi^2 x A_m(x) \sim \ln x + \ln 32 + \gamma - 2\left(1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2m-1}\right)$$
$$- (-)^m \frac{\pi^2}{4} \sqrt{\frac{\pi}{2x}} \cos\left(4x + \frac{\pi}{4}\right) + \cdots \qquad (90)$$
$$\gamma = \text{Euler's constant} = 0.5772 \ldots$$

As x approaches zero, S/I approaches $2P/Q$, which is to be expected because the frequency deviation of the unwanted carrier is represented by a pair of first order sidebands P/Q times as great as those on the wanted carrier. Averaging over the random carrier phase difference brings in a factor $\sqrt{2}$, and averaging over all frame phases accounts for another.

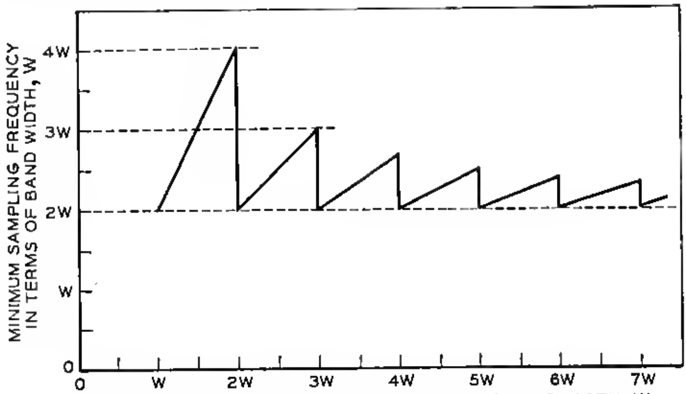


Fig. 37—Minimum sampling frequency for band of width W.

APPENDIX V

Sampling a band of frequencies displaced from zero

It is often necessary to transmit a signal band which does not extend all the way down to zero frequency. For example, a group of channels in FDM may be based on a set of carrier frequencies remote from zero. When we consider the application of pulse methods to transmit such a signal, the question of what sampling rate is needed immediately arises. A band extending from f_1 to $f_1 + W$ could of course be translated to the range 0 to W by standard modulation techniques, sampled at a rate $2W$, and restored to the original range by an inverse translation at the receiver. The frequency shifting apparatus required includes modulators, carrier generators, band separating filters, and possibly amplifiers to make up the inevitable losses.

A direct sampling process which avoids sbifting the band may therefore be preferred. A useful theorem for uniformly spaced samples is that the minimum sampling frequency is not in general twice the highest frequency in the band but is given by the formula:

$$f_r = 2W\left(1 + \frac{k}{m}\right), \tag{1}$$

where:

$$\begin{aligned} f_r &= \text{minimum sampling frequency} \\ W &= \text{width of band} \\ f_2 &= \text{highest frequency in band} \\ m &= \text{largest integer not exceeding } f_2/W \\ k &= \frac{f_2}{W} - m \end{aligned}$$

The value of k in (1) varies between zero and unity. When the band is located between adjacent multiples of W, we bave $k = 0$ and it follows that $f_r = 2W$ no matter how high the frequency range of the signal may be. As k increases from zero to unity the sampling rate increases from $2W$ to $2W$ $(1 + \frac{1}{m})$. The curve of minimum sampling rate versus the highest frequency in a band of constant width thus becomes a series of sawteeth of successively decreasing heigbt as shown in Fig. 37. The bighest sampling rate is required when $m = 1$ and k approaches unity. This is the case of a signal band lying between $W - \Delta f$ and $2W - \Delta f$ with Δf small. The sampling rate needed is $2(2W - \Delta f)$ which approacbes the value $4W$ as Δf approaches zero. Actually when $\Delta f = 0$, we change to the case of $m = 2$, $k = 0$, and $f_r = 2W$. The next maximum on the curve is $3W$, which is approached when f_2 nears $3W$. Tbe successive maxima decrease toward the limit $2W$ as f_2 increases. The sampling theorem contained in Eq. (1) may be verified from steady state modulation theory by noting tbat the first order sidebands on harmonics of $2W$ do not overlap the signal wben the equation is satisfied.